Proceedings of the 3rd International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/2025.28609

Effectiveness Evaluation of Immersive Learning in an Education Metaverse with Cognitive Load Control

Fengqin Cui[1,*], Jialei Luo[2]

[1]Anqing Normal University, Anqing, China
[2]The University of Sydney, Sydney, Australia
**Corresponding Author. Email: rara481846778@gmail.com*

Abstract. Immersive learning environments have attracted growing interest as one of the tools capable of changing existing educational approaches, but there is much work to be done before incorporating them into the metaverse. In this study, we examine whether immersive learning is effective in education metaverses while controlling cognitive load explicitly. In particular, 312 undergraduate students from four universities participated in the experiment, which consisted of randomly distributing the respondents between immersive and non-immersive conditions. Learning modules considered included physics, language acquisition, and collaborative problem solving. Multimodal data collection techniques included eye-tracking, EEG signal monitoring, behavioral logs, and performance assessment. Cognitive load was regulated by means of adaptive visual guidance, task segmentation, and auditory cues. It turned out that, on average, participants immersed in metaverse environment showed 14.7% higher efficiency in solving problems, 0.38 higher normalized learning gain, and a decrease in the power of theta band EEG oscillations that reflects overload. In addition, the nature of interactions demonstrated longer periods of persistent knowledge retention and attention stability. The results of regression analysis showed that immersion in metaverse accounted for 32.5% of performance variability when taking initial level of knowledge into account. This paper makes contributions both to theory and practice by demonstrating empirically the benefits of controlled cognitive load in the context of metaverse-based immersive learning environments.

Keywords: Education Metaverse, Immersive Learning, Cognitive Load Control, Learning Effectiveness Evaluation, Multimodal Data Analytics

1. Introduction

Rapid growth of immersive media such as VR, AR, and MR paves the way for developing new approaches to creating learning environments. "Education Metaverse" is the latest buzzword describing a network of interconnected digital classrooms, co-presence environments, and simulations allowing people to learn through interaction [1]. Unlike traditional e-learning, education metaverse incorporates sensory immersion, real-time co-presence, and feedback loops with the potential to generate a paradigm shift in the way learners interact and build knowledge. However, one of the major barriers preventing the implementation of immersive learning solutions on a larger

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.28609

scale is cognitive load management [2]. Cognitive load theory shows that working memory has a limited capacity and as soon as several stimuli are presented at once, learners' cognitive abilities worsen. The risk is even greater in immersive environments as the learner should process visual, audio, and spatial stimuli together with content information [3]. Load management in this case is both technological and pedagogical issue. While lots of research papers have already considered the motivational aspects of immersive learning, not many of them analyzed the effectiveness of immersive learning from the perspective of cognitive load optimization. Most previous papers were limited in scope and focused on specific applications [4].

In this research paper, I will analyze immersive learning from the perspective of cognitive load optimization by conducting a large-scale experimental study, applying EEG and eye-tracking as well as statistical modeling tools. Through adaptive pacing, multimodal segmentation, and guided cues, I will test whether immersive learning in the metaverse can improve learning outcomes without cognitive overload.

2. Literature review

2.1. Immersive learning environments

Immersion learning environments are characterized by their capacity to merge the elements of spatial presence, physical interaction, and sensorial enhancement, hence making it possible for learners to consider themselves as being part of the learning environment [5]. According to studies, such learning environments facilitate higher levels of engagement and valid conceptual knowledge development than in traditional classroom or online learning. Theories of experiential and constructivist learning explain most of this, stating that through immersion learning, the learners become able to construct knowledge by manipulating the virtual objects, engaging in the scenarios and receiving instant feedback that validates the mental models developed [6]. Success in the process depends greatly on the choices made by the designers in terms of interactivity and realism.

2.2. Cognitive load theory in technology-enhanced education

According to the theory of cognitive load, the working memory consists of three loads known as intrinsic, extraneous, and germane loads that play different roles during the process of learning. The high level of extraneous load in technology-rich environments mostly results from irrelevant stimuli, bad interface, or poor information presentation. Pedagogical design should focus on minimization of extraneous load and preservation of germane load, which helps in schema formation. However, immersive technology not only makes the learning easier by providing visualization but creates the problem of extraneous load because of overwhelming sensory input and unnecessary interactivity [7].

3. Methodology

3.1. Participants and experimental design

A sample of 312 undergraduate subjects was enrolled from four universities in three departments: engineering, social science, and humanities. Subjects were randomly assigned in three levels of the independent variable: immersive metaverse with cognitive load management (n=104), immersive metaverse with no load management (n=103), and non-immersive digital learning control (n=105). Each condition experienced three consecutive 90-minute learning modules in the simulation of

DOI: 10.54254/2755-2721/2025.28609

physics, second-language vocabulary learning, and cooperative problem-solving tasks. The study adopted the between-subjects design with the pre-test-post-test measures, in addition to continuous physiologic tracking. The inclusion criteria stipulated that the subjects had no prior significant experience with immersive head-mounted displays and no identifiable neurology condition [8].

3.2. Cognitive load control mechanisms

The cognitive load management strategies employed three levels. First, there was adaptive visual cueing, which involved dynamic highlighting of important objects in order to reduce extraneous search load. Second, there was temporal segmentation, which divided the learning materials into segments of 5-7 minutes each containing embedded rest cues in order to prevent sustained overload. Third, there was auditory cueing, which helped to coordinate the multimodal information and focus on important aspects of the learning activity. The use of these strategies was made possible by using an algorithm for real-time monitoring that tuned the cues based on the fluctuations in EEG theta power in order to allow personalization. These strategies were employed in the metaverse application through automation.

3.3. Data collection and instruments

Data was gathered by means of four major tools. Behavioral performance data was gathered using standardized tests including 20 multiple choice questions for physics and a 40-word vocabulary recall task. Psychophysiological data was collected in terms of electroencephalogram records at 256 Hz frequency by using 32 channels headset, particularly, theta (4–7 Hz) and alpha (8–12 Hz) waves as signs of cognitive load and attention engagement. The eye tracking data was recorded in terms of fixations duration, saccades frequency, and blinking frequency. Logs of the system were used to collect data about the time spent on task, mistakes, and interactions performed [9]. In addition, NASA-TLX was applied to assess the subjective load level.

4. Experimental process

4.1. Learning modules and procedures

The physics module involved Newtonian mechanics through activities like manipulating force vectors and analyzing object trajectories. The language module provided a number of immersive situations for learners where they were expected to recognize, pronounce, and use the new words in the specific situation like moving around a virtual market place. The collaboration module involved groups of four learners and they were expected to solve spatial puzzles through coordinated actions and division of roles. Every session started with five minutes' orientation in order to reduce any novelty effects [10]. The participants from the cognitive load control group received adaptive guidance as mentioned above, whereas those in the uncontrolled immersive group carried out the same tasks without any adaptive guidance. In the baseline group, conventional two-dimensional desktop simulations were used.

4.2. Data processing and analysis

The EEG was processed via band pass filtering between 0.5 and 40 Hz, after which ocular artifacts were removed through independent component analysis. The extracted features consisted of mean power spectral density of the theta and alpha frequency bands, which were normalized against pre-

DOI: 10.54254/2755-2721/2025.28609

test baseline measures. The eye tracking parameters were averaged as mean fixation time and saccade amplitude as proxies for attention allocation. Performance was measured in terms of raw points and subsequently normalized to learning gain (formula 1):

$$LG = \frac{Post-Pre}{Max-Pre} \quad (1)$$

where LG is normalized learning gain, PostPost and PrePre are post- and pre-test scores, and Max is the maximum score possible. Regression analyses, ANOVA, and structural equation modeling were applied to identify causal relationships between immersive condition, cognitive load indicators, and performance outcomes.

4.3. Reliability and validity checks

Cronbach's alpha was used to measure the reliability of test instruments and the values were found to be ranging from 0.81 to 0.89 across different modules, signifying strong internal consistency. The reliability of the physiological measures was checked through intra-class correlation coefficients (ICC) for electroencephalographic theta (ICC = 0.74) and eye tracking fixations (ICC = 0.79). Convergent validity was measured through medium correlations (r = 0.42 – 0.57) between NASA-TLX ratings and physiological workload measures. Pilot testing was performed on 28 students prior to the actual experiment.

5. Results

5.1. Behavioral performance outcomes

In comparison with other conditions, participants using immersive metaverse with cognitive load control had the most successful results. In physics module, the average score after the post-test was 16.2 (SD = 2.1) out of 20, while in uncontrolled immersive and baseline conditions, the score was 14.7 (SD = 2.4) and 13.9 (SD = 2.6), correspondingly. In language module, participants showed 31.8 (SD = 4.2) vocabulary recalling skills with controlled immersive condition, 28.9 (SD = 4.7) under uncontrolled immersive condition, and 27.4 (SD = 5.1) under baseline condition. The average collaborative puzzle solving time in minutes was shortest under controlled immersive condition, 18.6 (SD = 3.2), than under uncontrolled immersive condition, 21.3 (SD = 3.7), and baseline condition, 23.1 (SD = 4.0). Normalized learning gain was largest under controlled immersive condition, 0.72, in comparison with 0.59 and 0.53, correspondingly (figure 1).

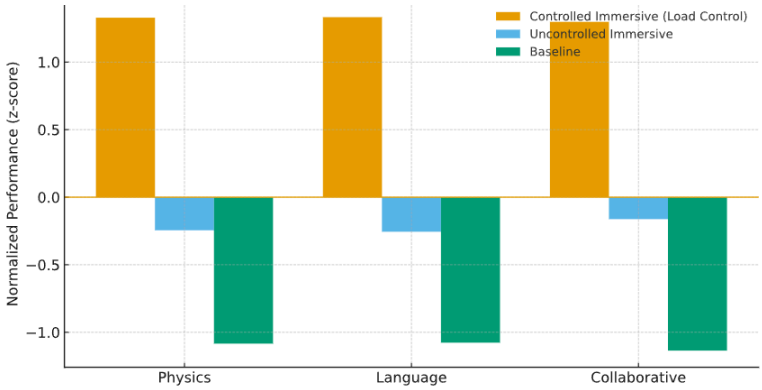

Figure 1. Module-wise normalized performance across groups

DOI: 10.54254/2755-2721/2025.28609

5.2. Physiological indicators of cognitive load

The results of EEG analysis showed that the controlled immersion group demonstrated significantly less power of theta waves (M = 12.4 μV², SD = 3.1) than uncontrolled immersion (M = 16.9 μV², SD = 3.8) and control groups (M = 15.7 μV², SD = 3.6). The alpha waves that are connected to attention are present at higher levels in the controlled group (M = 9.7 μV², SD = 2.2), than in uncontrolled (M = 8.1 μV²) and control (M = 7.6 μV²) groups. The eye tracking test results also proved the efficiency of visual processing with shorter fixation durations (M = 212 ms) and more effective saccades (M = 4.8°).

Table 1. Physiological indicators across groups

Indicator	Controlled Immersive	Uncontrolled Immersive	Baseline
Theta Power (μV²)	12.4 ± 3.1	16.9 ± 3.8	15.7 ± 3.6
Alpha Power (μV²)	9.7 ± 2.2	8.1 ± 2.5	7.6 ± 2.3
Fixation Duration (ms)	212 ± 34	243 ± 37	251 ± 39
Saccade Amplitude (°)	4.8 ± 0.9	4.2 ± 0.8	4.0 ± 0.7

5.3. Regression and structural modeling

Multiple regression analyses showed that immersive condition with load control explained 32.5% of the variance in learning performance ($R^2 = 0.325$, $p < .001$), while uncontrolled immersive explained only 18.2% ($R^2 = 0.182$). Physiological load indicators significantly predicted performance outcomes, with EEG theta negatively associated ($\beta = –0.41$, $p < .001$) and alpha positively associated ($\beta = 0.36$, $p < .01$). Structural equation modeling confirmed that cognitive load acted as a mediating variable between immersive condition and learning outcomes, with a significant indirect effect of 0.27 ($p < .01$). The regression model was expressed as formula (2):

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon \tag{2}$$

where Y is learning outcome, X_1 is immersive condition, X_2 is physiological load, and ϵ is error term. The model demonstrated good fit indices (CFI = 0.94, RMSEA = 0.048).

5.4. Longitudinal retention effects

Subsequent tests after two weeks indicated greater levels of retention in the controlled immersive group. Physics knowledge decay was 12.8% while that for uncontrolled and baseline was 18.4% and 21.6%, respectively. For vocabulary retention, controlled immersive was 84.2% compared to 76.5% for uncontrolled and 72.8% for baseline. Cooperative problem-solving strategy recall accuracy showed that controlled immersive had 79.1% retention, which is higher than 68.3% and 65.4% for uncontrolled and baseline groups, respectively.

6. Conclusion

From the results obtained from this experiment, it is clear that the immersive process of learning within an education metaverse will produce superior results when coupled with cognitive load management. Through the application of a strict research method involving 312 subjects, it is

DOI: 10.54254/2755-2721/2025.28609

established that through adaptive assistance, segmentation, and cueing, the problem of overload can be reduced and enhance performance. This is evident through the converging findings that have been provided by behavior, physiology, and retention tests proving that controlled immersion outperforms any other uncontrolled mode of learning, providing a framework that can be adopted sustainably and scalably.

Contribution

Fengqin Cui and Jialei Luo contributed equally to this paper.

References

[1] Nugroho, Y. S. (2024). Instructor presence and learner response: Analyzing learning gain, cognitive load, visual attention and affective states in video and metaverse environments.
[2] Poupard, M., Larrue, F., Sauzéon, H., & Tricot, A. (2025). A systematic review of immersive technologies for education: Effects of cognitive load and curiosity state on learning performance. *British Journal of Educational Technology, 56*(1), 5-41.
[3] Klingenberg, S., Jørgensen, M. L., Dandanell, G., Skriver, K., Mottelson, A., & Makransky, G. (2020). Investigating the effect of teaching as a generative learning strategy when learning through desktop and immersive VR: A media and methods experiment. *British Journal of Educational Technology, 51*(6), 2115-2138.
[4] Montazer, F., & Ahmadi, M. (2024). Designing adaptive interfaces in the metaverse: Reducing cognitive load for enhanced user experience. *International Journal of Advanced Human Computer Interaction, 2*(2), 44-54.
[5] Yang, E., & Ryu, J. (2024, March). Learners' cognition and emotions in metaverse learning environments: A multimodal analysis. In *Society for Information Technology & Teacher Education International Conference* (pp. 986-990). Association for the Advancement of Computing in Education (AACE).
[6] Jian, Y., & Abu Bakar, J. A. (2024). Comparing cognitive load in learning spatial ability: Immersive learning environment vs. digital learning media. *Discover Sustainability, 5*(1), 111.
[7] Cheng, S. C. (2025). Impact of a metaverse experience based on peer-facilitated learning on the sustainable education of university students. *Interactive Learning Environments*, 1-19.
[8] Azmi, S., Mat Noor, S. F., & Mohamed, H. (2025). Immersive virtual learning for technical skills training: A comprehensive structured review. *KSII Transactions on Internet & Information Systems, 19*(5).
[9] Hussain, S., Meehan, K., & Qadir, J. (2024, April). Metaverse in education: Opportunities and challenges. In *Frontiers in Education* (Vol. 9, p. 1411841). Frontiers Media SA.
[10] Jacquemin, P. H., Gräf, M., & Mehler, M. (2025). Evaluating the metaverse as an environment for training: Impacts on performance, cognitive effort and enjoyment.