A thesis submitted in partial satisfaction of the requirements
for the degree of Master of Computer Science and Engineering
in the Graduate School of the University of Aizu

The Role of Embodied Avatars and Generative AI in Self Learning VR Classroom



by

Lai Hui Shan

September 2025

© Copyright by Lai Hui Shan, September 2025

All Rights Reserved.

The thesis titled

The Role of Embodied Avatars and Generative AI in Self Learning VR Classroom

by

Lai Hui Shan

is reviewed and approved by:

Chief referee

Associate Professor Date

NASSANI Alaeddin

Senior Associate Professor Date

PYSHKIN Evgeny

Professor Date

WATANOBE Yutaka

THE UNIVERSITY OF AIZU

September 2025

Contents

Chapter 1 Introduction **1**
1.1 Background and Motivation . . . 1
1.2 Problem Statement . . . 2
1.3 Research Objective . . . 2
1.4 Proposed System . . . 2
1.5 Contributions . . . 3

Chapter 2 Related Work **5**
2.1 Embodied Avatars in Virtual Education . . . 5
2.2 Generative AI in Education . . . 9
2.3 Conversational AI in VR Environments . . . 14
2.4 Self-Directed Learning in VR Environment . . . 18
2.5 Mathematics Learning in VR Environment . . . 22
2.6 Interface Design and Usability in Educational Technologies . . . 26
2.7 Instructor Presence and Cognitive Load in Educational Media . . . 30

Chapter 3 System **34**
3.1 Original Vision and Design Evolution . . . 34
3.2 Prototype Overview . . . 39
3.3 Implementation . . . 43
3.3.1 Development Environment and Platform . . . 43
3.3.2 Video Learning Module . . . 45
3.3.3 AI Assistant Module . . . 47
3.3.4 Embodied Avatar . . . 50
3.3.5 User Interaction Logging . . . 53

Chapter 4 Evaluation **54**
4.1 Pilot Test . . . 54
4.2 User Study . . . 55
4.2.1 Task . . . 58
4.2.2 Conditions . . . 58
4.2.3 Measures Overview . . . 58
4.2.4 Adaptation from IPQ and NASA-TLX . . . 58
4.2.5 Avatar Perception Ratings . . . 60
4.3 User Study Results . . . 63
4.3.1 Participant Demographics (Pre-questions) . . . 63
4.3.2 IPQ and NASA-TLX Results . . . 65
4.3.3 Quiz Answers . . . 69
4.3.4 Ranking Results (post-questions) . . . 70

4.3.5 Avatar Perception Ratings 71

Chapter 5 Discussion **75**
5.1 Limitations 75
5.2 Future Work 76
5.3 User Feedback 76
5.3.1 Immersion and Classroom Realism 76
5.3.2 Lack of Interaction in the VR Setting 76
5.3.3 Perceived Potential and Limitations of the AI Assistant 77
5.3.4 Challenges for Non-native English Speakers 77
5.3.5 Avatar Design: Presence and Realism 77
5.3.6 Quiz Simplicity and Task Relevance 77
5.3.7 Note-taking Support in VR Environments 77
5.3.8 Implications for System Scalability 78

Chapter 6 Conclusion **79**

Chapter 7 Appendices **92**
7.1 Appendix A: Log Data 93
7.2 Appendix B: Questionnaire Survey 102
7.3 Appendix C: Quiz for Research 115
7.4 Appendix D: Feedback 118

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Prof. Alaeddin Nassani, for his constant and meticulous guidance throughout my research. His teaching not only helped me grasp the scientific approach to conducting research but also left a lasting impact that I will carry with me throughout my life.

I am also sincerely thankful to Prof. Julián Villegas for his support in statistical analysis and for helping edit the publications submitted to IEEE GEM and CHI. I also thank Prof. John Blake for his assistance in editing those publications.

I owe my deepest thanks to my family, without their unwavering support, I would not have been able to attend university, let alone graduate school. I am especially grateful to my classmate and beloved partner, He Runtong, whose care and encouragement gave me the courage to take on challenges I once thought were beyond me. I also wish to thank my friends for always being there for me.

Lastly, I want to thank myself, for never giving up and for holding on, no matter how hard things became.

Abstract

Background: Embodied avatars and virtual reality (VR) have been increasingly explored for their potential to enhance learner engagement and foster a sense of social presence in digital learning environments. At the same time, the growing use of generative AI in education has enabled more adaptive and interactive learning experiences. Combined, these technologies may offer valuable support for self-directed learning, especially in abstract domains such as calculus.

Objective: This study aims to investigate how the presence of embodied avatars and generative AI assistants influences learners' experiences in a self-directed, VR-based calculus learning environment, with particular attention to immersion, enjoyment, comprehension, and cognitive load.

Method: We developed a VR learning system that integrates generative AI, voice interaction, and embodied avatars. A within-subject user study ($n = 18$) was conducted to compare three instructional conditions: (1) Condition 1 – computer-based video learning; (2) Condition 2 – VR learning with a text-based AI panel; and (3) Condition 3 – VR learning with a speaking embodied AI avatar. Participants evaluated each condition using a 7-point Likert questionnaire, with items adapted from standardized instruments such as the NASA Task Load Index (NASA-TLX), the Igroup Presence Questionnaire (IPQ), and AttrakDiff.

Results: The results revealed that the embodied avatar condition led to statistically significant improvements in immersion and enjoyment compared to the baseline condition. However, no statistically significant differences were observed in comprehension and cognitive load.

Conclusion: The findings suggest that combining embodied avatars with generative AI may help enrich the emotional and interactive aspects of VR-based self-learning. This approach appears promising for enhancing engagement in abstract subjects such as mathematics. However, further research is needed to assess its long-term impact and broader applicability across diverse learning contexts.

Chapter 1

Introduction

In recent years, virtual reality and generative artificial intelligence have emerged as promising tools to support independent learning. This chapter introduces the motivation, problem context, and goals of the present research, which explores the intersection of these technologies in the context of mathematics education. It outlines the challenges of self-learning in advanced mathematics, the potential of immersive systems and embodied AI, and the rationale behind the system design. The research questions, hypotheses, and key contributions of this thesis are presented at the end of the chapter.

1.1 Background and Motivation

Mathematics is often perceived as abstract and difficult, presenting a persistent challenge for self-learners who lack access to real-time support or individualized guidance. Prior research indicates that students frequently struggle to sustain motivation and develop deep conceptual understanding under traditional instructional formats [1–3].

In response, immersive technologies, particularly virtual reality (VR), have emerged as promising tools for enhancing engagement and comprehension in educational contexts. Meta-analyses suggest that immersive VR can significantly improve learning outcomes and learner motivation [4, 5], especially in domains involving spatial or abstract content such as geometry and calculus [1, 6].

Concurrently, advances in generative artificial intelligence (AI), including large language models, have led to the development of intelligent conversational agents. These systems can provide real-time feedback, deliver personalize instructional content, and support self-regulated learning strategies [7–9]. Educational applications of generative AI now range from intelligent tutoring systems [10] to adaptive, dialogue-based learning support tools [11, 12].

A further development involves the use of embodied avatars, virtual agents with human-like appearance and behavior, in immersive learning environments. These avatars can replicate social cues such as eye gaze, gestures, and facial expressions, fostering more natural and engaging interactions than static or purely textual interfaces. Studies have shown that such embodiments can increase learner engagement and enhance perceived social presence [13–15].

Taken together, these developments highlight the potential for integrating immersive VR, embodied avatars, and generative AI to support more effective self-directed mathematics learning. However, a key question remains: how can such systems be designed to maximize educational benefit while preserving usability and learner motivation?

To address this, the present study investigates (1) whether the inclusion of visual embodiment in generative AI improves learner engagement and enjoyment, and (2) whether immersive VR learning environments provide advantages over traditional PC-based video learning in terms of immersion and perceived performance.

1.2 Problem Statement

Despite the promise of these technologies, empirical understanding remains limited regarding how embodied avatars, agents capable of speech, gestures, and facial expressions. However, their impact on learner experience in self-paced VR educational contexts remains underexplored. While prior work has explored static AI panels and text-based assistants [10, 11], few systems have integrated both *visual embodiment* and *generative conversational AI* within immersive learning environments. This leads to the central research question: To what extent can embodied AI avatars improve user immersion, engagement, and learning perception compared to traditional or non-embodied approaches?

1.3 Research Objective

This study aims to systematically investigate the impact of embodied avatars and generative AI on self-directed learning in VR-based mathematics education. Specifically, it compares three instructional conditions:

- Condition 1: PC-based video learning (baseline).
- Condition 2: VR learning system with a text-based AI assistant.
- Condition 3: VR learning system with a speaking, visually embodied AI avatar. avatar.

A user study involving 18 participants was conducted to evaluate differences across these conditions in terms of immersion, enjoyment, cognitive load, and perceived comprehension.

1.4 Proposed System

To support this investigation, we developed a VR-based learning system that integrates a voice-enabled generative AI assistant. The system employs OpenAI's GPT-4o for response generation, Whisper for speech recognition, and Amazon Polly for speech synthesis. In the embodied condition, the AI assistant is represented by a 3D avatar created using Ready Player Me, with lip-sync animations driven by Oculus LipSync. This enables more realistic and responsive interactions between users and the AI assistant. Participants interact with the system via voice input, receiving either textual responses (Condition 2) or spoken responses with full embodiment (Condition 3), depending on the experimental setup.

To provide an overview of the structure and logical flow of this thesis, Figure 1.1 presents a mindmap that summarizes the organization of all chapters and their major

subsections. This visual serves as a conceptual roadmap from system design to evaluation and discussion.

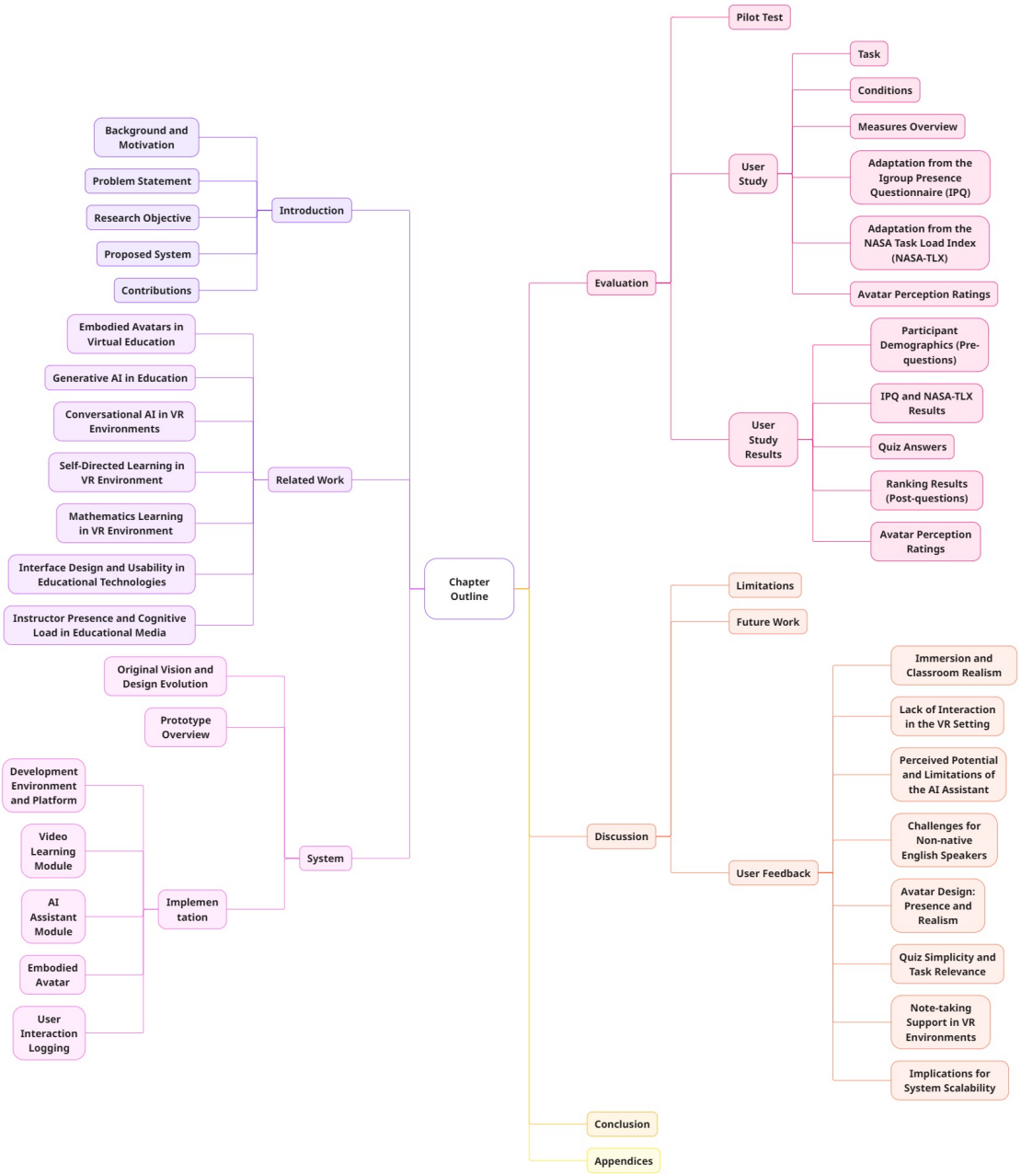


Figure 1.1: Chapter structure overview of this thesis. The mindmap visualizes all major chapters and subsections.

1.5 Contributions

This work aims to address the following research questions:

- RQ 1: How does adding an embodied avatar to a generative AI assistant affect learners' engagement and enjoyment in a self-learning VR classroom?

- RQ 2: How does learning in a VR-based classroom compare to traditional PC-based video learning in terms of learner immersion, enjoyment, and quiz performance?

The hypotheses are:

- H1: Participants will report higher engagement and enjoyment in the condition with an embodied AI avatar (C3) compared to the non-embodied AI panel (C2).
- H2: Participants will report higher immersion and enjoyment, and achieve better quiz performance in each VR-based condition (C2 and C3) compared to the PC-based condition (C1).

This thesis presents three main contributions:

1. A novel VR-based self-learning system that combines AI-driven speech interaction with embodied avatars.
2. A within-subject user study design comparing different learning modalities and their effects on subjective experience.
3. Empirical evidence highlighting the benefits and limitations of embodied AI in VR education, with implications for future research on intelligent tutoring systems.

Chapter 2

Related Work

This chapter reviews prior studies that inform the design and evaluation of the VR-based AI learning system developed in this thesis. It covers seven key areas: embodied avatars in virtual education, the use of generative AI in learning contexts, conversational AI in VR environments, self-directed learning strategies supported by VR, mathematics learning in immersive settings, the role of interface simplicity, and instructor presence in educational technologies. Together, these areas provide the conceptual foundation and practical context for understanding how immersive systems can support effective, engaging, and personalized mathematics learning.

2.1 Embodied Avatars in Virtual Education

Embodied avatars have gained increasing attention in educational research for their potential to improve engagement, comprehension, and learning outcomes. The concept is grounded in embodied cognition theory, which emphasizes that learning arises from the interaction of the brain, body, and environment [16]. Castro-Alonso et al. [16] identify six research avenues supporting this view: physical activity, generative learning, offloaded cognition, specialized processor, signaling, and social cognition. These avenues highlight how both producing and observing human movements may support learning, suggesting that educational systems could benefit from incorporating bodily cues and actions into instructional design.

In immersive virtual environments, avatars with bodily representation allow learners to interact in ways that feel more natural and socially connected. Schroeder and Craig [14] describe virtual humans as embodied agents with human-like appearances and outline promising directions for future research in this field. They emphasize the diversity of methods and applications in virtual human studies and suggest significant potential for improving educational experiences.

Empirical evidence supports the effectiveness of embodied avatars, although results vary depending on design and context. Mizuho et al. [15] introduced a learning method using multiple self-avatars in VR to teach sign language. While learners with varied avatars initially performed worse, they retained significantly more information after one week, showing that diverse embodied contexts can improve long-term memory retention. The study also found a positive correlation between perceived avatar embodiment and learning outcomes, demonstrating the value of embodied interaction in memory-related tasks.

Davis et al. [13] conducted a meta-analysis on embodied pedagogical agents and

found that such agents significantly improve learning. However, effectiveness depends on testing formats and agent design. Their findings suggest that evaluating pedagogical agents requires nuanced assessment strategies that go beyond traditional test categories.

In the context of programming education, Parmar et al. [17] found that middle school students using immersive VR with self-avatars outperformed those using non-embodied or desktop-based VR. The embodied VR group also reported more positive attitudes toward computer science and higher levels of presence and confidence, indicating that embodiment enhances not only performance but also learner perception.

Studies in specialized fields have also explored embodied avatars. Cakmak et al. [18] developed a human muscular arm avatar to teach hand and forearm anatomy. Most students found the tool engaging and useful, and some explicitly noted an embodied learning experience. The ability to visualize abstract anatomy through avatar-based embodiment was perceived as advantageous over static models.

Embodied VR has also been shown to support real-world motor learning. Haar et al. [19] built a system to simulate billiards in VR with physical components, finding that motor learning patterns in VR closely matched those in the physical world. This suggests that embodied VR environments can facilitate skill transfer beyond virtual settings.

Despite these promising results, some studies have reported mixed findings. Fitton et al. [20] reported no significant differences in learning outcomes, motivation, or engagement when using embodied agents in PowerPoint-based e-learning scenarios. Their findings raise important questions about the contexts in which embodiment is beneficial and call for further research to optimize design in remote education.

Learner characteristics may also moderate the impact of embodiment. Chiquet et al. [21] found that individuals with high absorption tendencies learned more effectively when teaching an agent in immersive VR, while those with low absorption benefited more from imagining helping a peer. These findings suggest the need to tailor VR learning systems to individual differences.

In the field of augmented reality, Genay et al. [22] reviewed virtual self-avatar embodiment and proposed a continuum to categorize embodiment experiences. They discussed implementation challenges and potential applications in education, communication, and healthcare, calling for more research to better understand how avatar embodiment functions in AR contexts.

Taken together, these studies suggest that embodied avatars may enhance learning through mechanisms such as increased memory retention, deeper engagement, and stronger spatial presence. However, their effectiveness depends on thoughtful design, task relevance, and learner characteristics. Further research is needed to identify when and how embodiment should be implemented in educational technologies.

To synthesize the scope and diversity of existing research, we constructed a thematic mind map based on ten representative studies (see Figure 2.1). This visual classification highlights key clusters such as embodied cognition theory, empirical outcomes, application domains, benefits, and methodological challenges. It offers an intuitive overview of how prior work is distributed across different conceptual dimensions.

Table 2.1 further provides a comparative summary of these studies, outlining their key findings and identifying the remaining gaps in the literature.

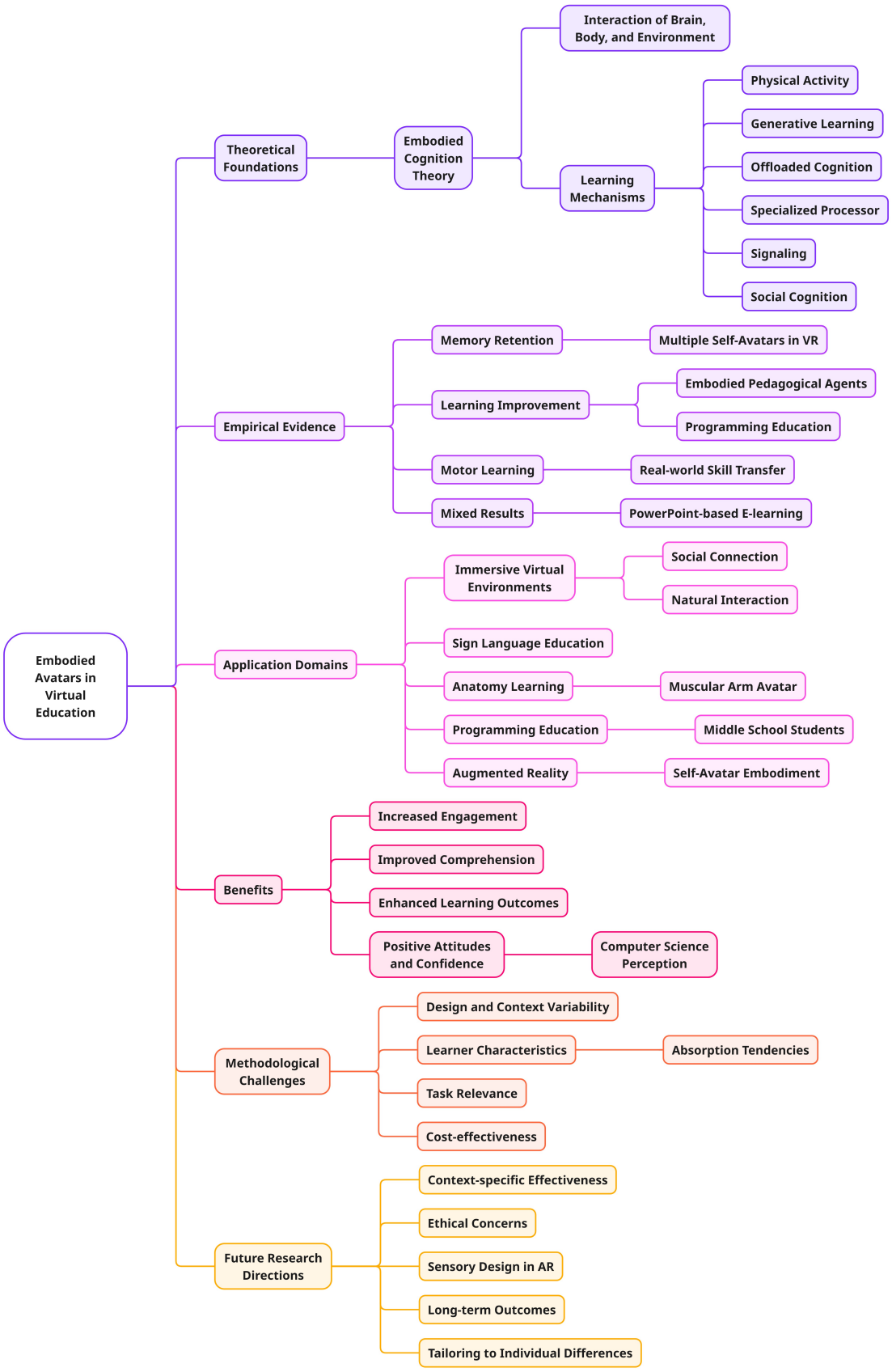


Figure 2.1: Mind map of existing studies on embodied avatars in virtual education. This diagram organizes prior research into thematic branches such as theoretical foundations, empirical evidence, benefits, applications, challenges, and future research directions.

Table 2.1: Summary and Identified Gaps in Related Work on Embodied Avatars

Study	**Summary**	**Gap**
Castro-Alonso et al. [16]	Summarizes six embodied cognition mechanisms that support learning.	Future work should test which factors most influence effectiveness in each pathway (e.g., gesture vs. observation).
Schroeder and Craig [14]	Reviews virtual humans as instructional agents and research directions.	Abstract lacks specific gaps; future studies should define concrete open questions.
Mizuho et al. [15]	Using multiple self-avatars in VR enhances memory retention over time.	Lacks clarity on which contextual elements drive the effect; needs testing across tasks, populations, and complex environments.
Davis et al. [13]	Meta-analysis shows embodied agents improve learning depending on design and content type.	More research needed on labeling bias, agent design features, and context-specific effectiveness.
Parmar et al. [17]	Self-avatars in immersive VR improve programming performance and attitudes.	No clear gaps identified; limited detail on mechanisms behind improvement.
Cakmak et al. [18]	Muscular arm avatar helped students learn anatomy through body ownership.	More comparisons with traditional methods are needed to validate learning advantages.
Haar et al. [19]	VR motor training showed similar learning patterns to real-world tasks.	Needs deeper analysis of transfer to real life and effects of avatar absence.
Fitton et al. [20]	Embodied agents in PowerPoint-based e-learning showed no learning benefit.	Questions remain about when embodiment adds value and how to improve cost-effectiveness.
Chiquet et al. [21]	VR teaching is more effective for learners with high absorption traits.	More work needed on user traits, plausibility, novelty effects, and long-term outcomes.
Genay et al. [22]	Reviews avatar embodiment in AR and proposes a categorization framework.	Calls for more education-focused research, including sensory design and ethical concerns.

2.2 Generative AI in Education

Generative artificial intelligence (AI) has rapidly emerged as a transformative force in education, offering opportunities for personalized learning, automated content creation, and adaptive instructional support. Building on advances in large language models and multimodal generation, researchers and educators are exploring how these technologies can be integrated into classrooms, virtual learning environments, and intelligent tutoring systems to enhance teaching and learning outcomes.

In mathematics and STEM education, generative AI has demonstrated the ability to support both procedural understanding and conceptual insight. Drori et al. [23] presented a neural network capable of solving, explaining, and generating university-level mathematics problems with near human-level performance. Their system, trained using few-shot learning and text-code hybrid models, demonstrated the potential for automating course content development and assessment. Similarly, Davies et al. [24] proposed a machine learning framework for discovering mathematical conjectures and guiding theorem proving, highlighting how AI can augment creative mathematical reasoning and research.

AI-driven tools have also been applied to educational content generation and engagement. Chen and Wu [25] designed a generative AI system that creates Tang poetry videos through text-to-image and text-to-speech synthesis. Their system improved comprehension and reduced cognitive load, demonstrating how multimodal instructional materials can enhance learning experiences. In higher education, Pesovski et al. [8] developed a generative system that produced multiple themed versions of course content and quizzes. Students found these personalized variants engaging, and the system appeared to increase study motivation and time, especially among learners with weaker prior knowledge.

Within classroom contexts, Sajja et al. [7] introduced the AI-Enabled Intelligent Assistant (AIIA), which integrates with learning management systems to deliver personalized feedback, generate flashcards and quizzes, and support adaptive learning. The framework aims to reduce cognitive overload and improve student engagement, although real-world classroom validation remains ongoing. Qawaqneh et al. [10] evaluated an AI-based virtual mathematics laboratory for middle school students and found it significantly improved learning motivation compared to both traditional and 3D visual systems, reinforcing the motivational value of AI-assisted environments.

Beyond specific systems and applications, broader reviews of AI integration into educational practices show a clear progression in technological sophistication. Chen et al. [12] traced the evolution of AI in education from early computer-based technologies to modern web-enabled intelligent systems and humanoid robots. Their study highlights how AI tools have been used to support administrative efficiency (such as automated grading), personalized curriculum delivery, and instructional interaction. By customizing content to individual learner needs and reducing instructors' repetitive tasks, AI systems have improved engagement and teaching quality across a variety of settings. The authors emphasize that AI's expanding role in education requires ongoing pedagogical reflection to ensure that technological advancement translates into meaningful learning gains.

In professional education domains such as engineering and medicine, generative AI offers innovative solutions to persistent instructional challenges. Menekse [26] discussed how AI can help overcome limitations in lecture-driven engineering education

by enabling the generation of diagrams, summaries, and instructional media tailored to learner needs. Boscardin et al. [27] explored the role of generative AI in medical education, highlighting both the instructional opportunities (e.g., simulation, assessment) and associated risks (e.g., plagiarism, misinformation). They advocated for the development of AI literacy among medical educators to responsibly integrate these technologies into training environments.

K-12 education is also beginning to benefit from generative AI. Kong and Yang [28] proposed a human-centered design framework for integrating AI into primary education, emphasizing personalized feedback, engagement, and self-regulated learning. Their teacher development program showed increased teacher confidence in designing AI-supported curricula and improvements in students' attention and satisfaction. The framework positions educators not just as content deliverers but as facilitators and storytellers, guiding AI-enhanced learning experiences.

Despite the potential, scholars have raised important ethical and pedagogical concerns. Kohnke et al. [29] and Dergaa et al. [30] caution that tools like ChatGPT may foster dependency, bias, or misuse if deployed without oversight. Their studies emphasize the importance of critical thinking, digital literacy, and ethical awareness in AI-integrated learning environments. Similarly, Mittal et al. [31] and Cooper [32] warn against uncritical acceptance of AI-generated content and call for transparent, accountable practices when using generative AI for educational tasks such as writing or assessment design.

Student perspectives also reflect both optimism and concern. Chan and Hu [33] surveyed university students and found generally positive attitudes toward generative AI in higher education, particularly for writing assistance and personalized support. However, students also expressed anxiety about accuracy, ethical use, and long-term career impact. Mejeh and Rehm [34] highlighted the importance of aligning AI tools with adaptive learning strategies and self-regulated learning goals. Their year-long co-design study found that natural language processing (NLP), when integrated with adaptive learning technologies, can provide meaningful and individualized support. However, further refinement is needed to ensure pedagogical effectiveness.

Bahroun et al. [9] conducted a comprehensive review of 207 studies on generative AI in education, noting a surge in research interest, especially around tools like ChatGPT. They identified major themes including AI in assessment, content creation, and intelligent tutoring, while also emphasizing the need for interdisciplinary collaboration and responsible model design. The authors advocate for future studies to examine long-term learning impacts and the role of AI in curriculum development.

Overall, the integration of generative AI in education offers numerous advantages, including personalized support, efficient content generation, and expanded access to learning resources. However, realizing this potential requires careful attention to design, ethical safeguards, educator training, and empirical validation. As AI continues to evolve, its role in shaping the future of learning will depend on how educators, researchers, and policymakers collaboratively navigate its promises and limitations.

To provide a structured overview of representative studies in this domain, we summarize the main findings and research gaps in Table 2.2. This tabular synthesis highlights diverse applications—from STEM instruction and adaptive learning systems to ethical concerns and student perspectives—while also identifying unresolved challenges that warrant further exploration.

Furthermore, Figure 2.2 presents a thematic mind map that visualizes the concep-

tual landscape of generative AI in education. The diagram categorizes existing research into core areas such as educational domains, application types, pedagogical benefits, challenges, and future directions. This classification serves as a visual guide to understanding the scope and complexity of current discourse surrounding generative AI technologies in learning contexts.

Table 2.2: Summary and Identified Gaps in Generative AI Research in Education

Study	**Summary**	**Gap**
Drori et al. [23]	A neural network is demonstrated to solve, explain, and generate math problems using few-shot learning and program synthesis.	Neural networks were previously believed unable to do symbolic math; more evaluation needed across mathematical domains.
Davies et al. [24]	Proposes a framework where machine learning aids mathematical discovery and conjecture generation.	Effectiveness depends on large datasets and domain-specific models; generalizability not fully tested.
Chen and Wu [25]	Presents a generative AI system that creates Tang poetry videos via text-to-image and speech synthesis.	More research needed on combining multimodal generation for learning beyond poetry or humanities.
Pesovski et al. [8]	Proposes a system that generates multiple versions of course materials and quizzes to enhance engagement.	Lack of empirical classroom studies; needs further testing with varied learner backgrounds.
Sajja et al. [7]	Introduces an AI-based assistant that provides personalized quizzes and feedback via LMS integration.	Real-world validation in live classrooms is limited; longitudinal effects on learning unknown.
Qawaqneh et al. [10]	Investigated AI-based virtual math labs for middle school students and found improved motivation and learning outcomes.	More research needed on technical infrastructure and long-term learning gains.
Chen et al. [12]	Reviews the evolution of AI in education from early tools to modern intelligent systems and robots.	Lack of critical discussion on pedagogical alignment and long-term impacts.
Menekse [26]	Discusses how AI-generated visuals and summaries can support personalized engineering education.	No specific gap stated; future studies should explore practical classroom implementation.
Boscardin et al. [27]	Examines both potential benefits and risks of generative AI in medical education, including simulation and misinformation.	Calls for AI literacy among educators and safeguards against bias and misuse.

Study	Summary	Gap
Kong and Yang [28]	Proposes a human-centered framework for AI-supported K–12 learning that promotes teacher involvement and self-regulation.	Need for more empirical studies on teacher-student-AI dynamics and long-term learning.
Kohnke et al. [29]	Explores ChatGPT's usefulness in language learning, including writing support and engagement.	Questions remain about dependency, accuracy, and fostering critical thinking.
Dergaa et al. [30]	Assesses ChatGPT's potential and risks in academic writing, including productivity and plagiarism.	Lacks a structured ethical framework for responsible use in education.
Mittal et al. [31]	Reviews generative AI's role in writing, assessment, and creativity in education.	Privacy, transparency, and assessment fairness are key unresolved concerns.
Cooper [32]	Evaluates ChatGPT's accuracy and appropriateness in science education questions.	Absence of studies validating ChatGPT content against curriculum standards.
Chan and Hu [33]	Surveys student views on generative AI, finding benefits for support but concerns over ethics and career impact.	Broader sampling and longitudinal tracking needed to generalize insights.
Mejeh and Rehm [34]	Explores how NLP-based adaptive learning systems align with self-regulated learning goals.	Lack of pedagogical framework and refinement for sustained educational use.
Bahroun et al. [9]	Comprehensive review of 207 studies on generative AI in education, covering assessment, content creation, and tutoring.	More longitudinal studies and interdisciplinary approaches needed to assess impact.



Figure 2.2: Thematic mind map of generative AI in education, illustrating key domains, application areas, pedagogical benefits, implementation challenges, and future directions.

2.3 Conversational AI in VR Environments

Conversational AI has been increasingly integrated into virtual reality (VR) systems across domains such as education, healthcare, and professional training. These integrations aim to improve user engagement, support skill development, and enhance learning outcomes.

In educational settings, various studies have demonstrated the value of combining conversational agents with immersive VR experiences. Divekar et al. [35] developed the Cognitive Immersive Language Learning Environment (CILLE), which merges AI and extended reality to promote Chinese vocabulary and conversation skills. Their seven-week study showed significant gains in language comprehension and learner engagement. Similarly, Xie et al. [36] used a motivational conversational agent within an immersive environment to improve collaborative learning. The agent, designed with motivational interviewing techniques, led to measurable improvements in student engagement. Wang et al. [37] integrated ChatGPT into a VR-based AIoT training program and found that personalized feedback improved students' reflective thinking, learning motivation, and hands-on performance. Taylor et al. [38] demonstrated the use of conversational AI in digital twin laboratories, where it provided interactive guidance and supported scientific exploration in chemical education.

Conversational AI has also been applied to healthcare training and support. Liaw et al. [39] designed an AI-enabled VR simulation that improved nursing students' interprofessional communication skills. While students rated the system positively for usability and learning effectiveness, they noted that the AI doctor's expressiveness could be improved. Sardesai et al. [40] created a virtual patient avatar for anaesthesia training using a no-code platform. The tool was rated highly for realism and usability by professional anaesthetists, although concerns were raised about the potential for AI-generated misinformation. In a different context, Garzotto et al. [41] explored the use of conversational agents within VR and AR environments to support individuals with autism in performing daily tasks. Their study with 17 participants showed that the combination of speech-based agents and cross-reality training helped improve autonomy and real-world skill application.

Training in industrial settings has also benefited from AI-VR integration. Hussain et al. [42] developed a VR safety training system for migrant construction workers using ChatGPT as an embedded conversational instructor. The study, involving participants from five countries, reported a 23% increase in knowledge scores after training. This result suggests that conversational AI can effectively support safety training across diverse linguistic and educational backgrounds.

Conversational agents are also being explored for health promotion and consumer interaction. Luo et al. [43] conducted a systematic review on promoting physical activity using conversational agents. While results showed moderate usability and effectiveness, challenges included repetitive content, high dropout rates, and privacy concerns. Bergner et al. [44] investigated conversational AI in brand interaction, demonstrating that dialogue features such as turn-taking and grounding enhanced perceived humanness and fostered stronger consumer-brand relationships.

In language education, Ji et al. [11] reviewed the role of conversational agents in supporting teaching activities. While most studies focused on AI as a learning tool, few demonstrated meaningful collaboration between teachers and AI systems. The authors advocate for future systems that integrate conversational AI to amplify teacher capacity

and orchestrate classroom activities more effectively.

A broader overview of AI in VR is provided by Ribeiro De Oliveira et al. [45], who conducted a systematic literature review on AI-enhanced VR applications across domains including education, agriculture, and healthcare. They concluded that AI improves algorithmic precision and user interaction but noted a lack of research on long-term reliability and ethical concerns such as privacy.

Together, these studies highlight the growing potential of conversational AI in VR. While current applications show benefits across diverse contexts, challenges remain regarding system reliability, content quality, expressiveness, and ethical considerations. Continued research is needed to refine these systems and ensure their effectiveness in real-world learning and communication environments.

To synthesize current findings and guide future research, we summarize the key studies in Table 2.3, highlighting their main contributions and remaining gaps in implementation, generalization, and evaluation.

To further organize the conceptual space of this literature, Figure 2.3 presents a thematic mind map. This visual classification outlines core domains—education, healthcare, industrial training, and more—alongside implementation challenges and identified gaps. It helps visualize how conversational AI in VR is being explored across multiple sectors while revealing cross-cutting concerns such as ethical design, personalization, and educational alignment.

Table 2.3: Summary and Identified Gaps in Conversational AI Research in VR

Study	**Summary**	**Gap**
Xie et al. [36]	Designed a motivational conversational agent using motivational interviewing in VR to support collaborative learning.	Needs better integration of dialog strategies with educational goals and improved assessment of engagement over time.
Wang et al. [37]	Integrated ChatGPT into a VR AIoT course to provide adaptive feedback and boost reflective thinking and motivation.	Current system lacks a structured instructional model; more real-world studies are needed.
Taylor et al. [38]	Demonstrated conversational AI for guiding learners in VR-based digital chemistry labs.	Lacks operational models for integrating AI with real-time lab environments and diverse learner needs.
Liaw et al. [39]	Developed a VR simulation with conversational AI to train nursing students in interprofessional communication.	Expressiveness of the AI agent needs enhancement; lacks evaluation across varied clinical contexts.
Sardesai et al. [40]	Created a VR avatar for anesthesia training using a no-code platform, rated high in realism and usability.	Risk of misinformation remains a concern; platform needs expert oversight and verification mechanisms.
Garzotto et al. [41]	Used conversational agents in VR and AR to support autistic individuals in daily tasks via cross-reality training.	More work needed on personalization and long-term outcomes in assistive environments.

Study	Summary	Gap
Hussain et al. [42]	Developed a VR safety training program with ChatGPT for migrant construction workers; improved knowledge and engagement.	Needs broader validation across different job types and AI language challenges.
Luo et al. [43]	Reviewed conversational agents for promoting physical activity; found moderate effectiveness with engagement limitations.	Challenges include repetitive content, privacy concerns, and high dropout rates.
Bergner et al. [44]	Studied brand interaction using conversational AI; turn-taking and grounding increased perceived humanness.	Further studies needed on how conversation design affects trust in commercial vs. educational settings.
Ji et al. [11]	Reviewed conversational agents in language education; found underuse in teacher-AI collaboration.	Future systems should enhance collaboration between teachers and AI to support classroom orchestration.
Ribeiro De Oliveira et al. [45]	Reviewed AI-enhanced VR across domains; found increased interaction and accuracy but ethical issues persist.	Long-term reliability, privacy risks, and ethical design need more research attention.
Divekar et al. [35]	Developed a cognitive immersive VR environment for Chinese language learning; improved comprehension and engagement.	Needs more studies on long-term retention, transferability, and conversational scalability.

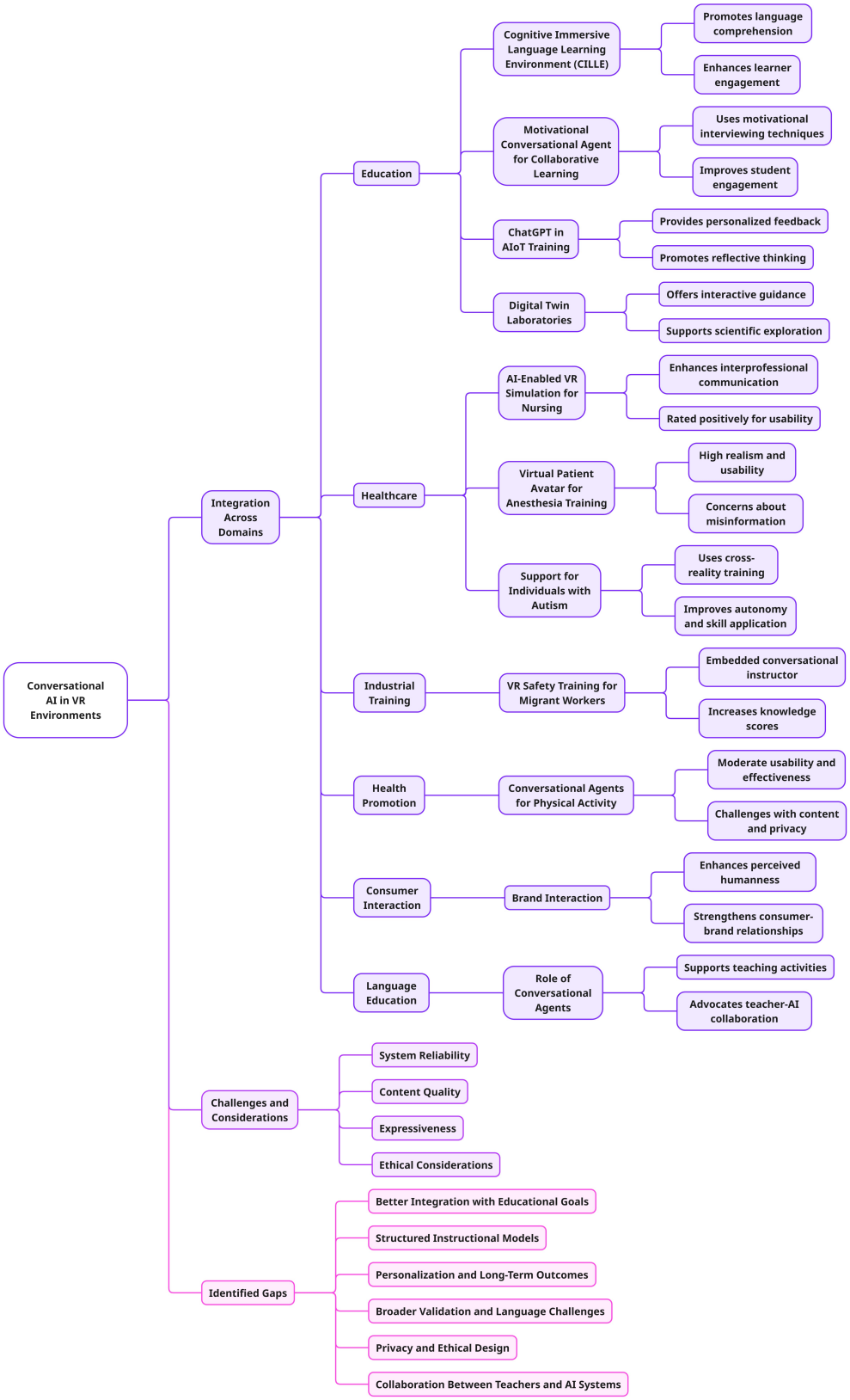


Figure 2.3: Thematic mind map of conversational AI in VR environments. Domains include education, healthcare, industrial training, health promotion, and language education. Key considerations involve reliability, content design, ethical use, and identified research gaps such as personalization, teacher-AI collaboration, and long-term outcomes.

2.4 Self-Directed Learning in VR Environment

Self-regulated learning (SRL) is essential for promoting autonomy and deep understanding in education, and virtual reality (VR) environments have shown increasing potential to support these processes. Unlike traditional instructional settings, VR can offer learners immersive, interactive, and personalized experiences that facilitate SRL strategies such as goal setting, self-monitoring, strategic planning, and reflection.

Recent studies have explored how various VR features influence self-learning outcomes. Sakdavong et al. [46] examined the role of immersion and learner control in a VR museum for art education. While immersive display modes alone did not significantly affect motivation or learning outcomes, the ability to actively control the learning process enhanced performance. Both immersion and control were found to partially support SRL, indicating that the value of VR may depend more on how learners interact with the content than on immersion alone.

In the domain of language learning, Chen and Hsu [47] evaluated a VR mobile game designed to improve English proficiency among university students. They found that game-based interaction, combined with features promoting presence and flow, enhanced learners' intrinsic motivation and self-efficacy. Immersion and absorption significantly influenced students' self-regulation, and the reciprocal relationship between self-efficacy and SRL behavior suggested a reinforcing learning cycle. These results support the use of interactive, game-like VR platforms for promoting both engagement and independent learning strategies.

Jiang et al. [48] explored how immersive virtual environments (IVEs) can foster SRL among middle school students learning scientific inquiry. Their study revealed that IVEs encouraged learners to develop and refine note-taking, reviewing, and problem-solving strategies over time. Students learned to navigate multiple sources of information, track their progress, and adapt their learning behaviors, all of which are critical SRL skills in complex learning environments.

The relationship between cognitive load and SRL has also been studied in immersive VR contexts. Sobocinski et al. [49] collected behavioral, verbal, and physiological data to explore how learners monitored their understanding during a VR-based game on exponential functions. The study found that heart rate variability, used as a proxy for cognitive load, predicted episodes of metacognitive monitoring. Additionally, physical motion in the VR environment often coincided with reflective thinking. These findings highlight the embodied nature of cognition in VR and suggest that real-time physiological measures may be used to detect and support SRL dynamically.

Beyond individual studies, Neo et al. [50] conducted a systematic review on the integration of SRL strategies in AI-driven recommender systems within virtual learning environments. The review emphasized that such systems can guide learners through personalized pathways based on their goals and behaviors, enhancing academic outcomes while preserving learner agency. The authors highlighted the potential of combining VR with adaptive systems to deliver timely, personalized support aligned with SRL principles.

Addressing support for lower-performing learners, Wang et al. [51] implemented a VR course on electronic circuits that incorporated learning aids such as hints, feedback, and guidance. A quasi-experimental study showed that these aids significantly improved practical skills, cognitive performance, and SRL ability. The research demonstrates that well-designed scaffolding in VR can help bridge performance gaps for learn-

ers who may struggle with independent regulation.

Other studies have developed personalized VR environments to target specific domains. Wu et al. [52] designed a spherical video-based VR system for landscape architecture students, addressing the lack of real-world context during pandemic-related remote learning. The system improved learning outcomes, attitudes, and SRL behavior compared to traditional instruction. However, it required time for learners to build confidence and self-efficacy, suggesting the need for gradual onboarding and support.

Yang et al. [53] created a VR-based cardiopulmonary resuscitation (CPR) training system that combined video instruction, feedback mechanisms, and game-like self-practice modules. In tests with students unfamiliar with CPR, 85% completed intensive practice sessions, and significant gains were observed in skill acquisition and interest. The system demonstrated how VR can enable accessible, low-cost, and self-directed skill training.

To support the assessment of SRL skills, Koçdar et al. [54] developed a validated scale for use in open and distance learning environments. The scale consists of five components—goal setting, help seeking, self-study strategies, environmental management, and effort regulation—and showed strong psychometric properties. This instrument fills a gap in measuring SRL in self-paced VR and online learning contexts, offering a tool for researchers and educators alike.

In addition to pedagogical design, the technological backbone of VR learning is also evolving. Wang et al. [55] reviewed applications of deep learning in VR content creation. They described how generative adversarial networks and convolutional networks have been used to support adaptive environments, enabling dynamic customization of content, environments, and learner feedback loops. Such intelligent VR systems can facilitate SRL by automatically adjusting learning pathways based on user behavior and progress.

Villena-Taranilla et al. [4] provided broader empirical support for the use of immersive VR in early education. Their meta-analysis of 21 studies found that VR-based learning significantly improved outcomes across various knowledge domains for K-6 students. Immersive systems had the highest effectiveness, and shorter sessions tended to produce stronger effects. The results highlight the suitability of VR for self-paced, exploratory learning in foundational education.

Overall, the literature supports the notion that VR, when designed with SRL principles in mind, can effectively promote autonomous learning. Critical design elements include learner control, personalized feedback, motivational immersion, and adaptive system features. At the same time, real-time measurement tools and targeted scaffolding can help address challenges such as cognitive overload and motivational fatigue. Future work should continue exploring how to balance autonomy and support in VR environments to cultivate lifelong self-learners.

To further illustrate the conceptual structure of self-directed learning in VR, Figure 2.4 presents a thematic mind map synthesized from the key studies summarized above. It highlights major themes including cognitive load, personalized VR environments, adaptive systems, assessment methods, and technological enablers. This visualization supports a comprehensive understanding of how VR features interact with SRL strategies and where future research can be directed.

Table 2.4: Summary and Identified Gaps in Conversational AI Research in VR

Study	**Summary**	**Gap**
Sakdavong et al. [46]	Examined how immersion and control in a VR art museum affect motivation and SRL. Found learner control positively influenced performance.	Further work is needed to isolate the effect of immersion and understand long-term self-regulation impact.
Chen and Hsu [47]	Evaluated a VR English-learning game promoting flow and motivation; found mutual reinforcement between self-efficacy and SRL.	Requires deeper study of game mechanics and presence in shaping SRL behavior across disciplines.
Jiang et al. [48]	Explored SRL in immersive science inquiry tasks for middle schoolers using IVEs.	SRL strategy development needs finer analysis over time, especially in younger learners.
Sobocinski et al. [49]	Used physiological and motion data to detect cognitive load and metacognition during a VR math game.	Embodied metrics of SRL are under-explored and require standardization.
Neo et al. [50]	Reviewed SRL-driven AI recommender systems in virtual learning; emphasized adaptive support and learner agency.	Lacks integration with immersive systems like VR; hybrid systems remain untested.
Wang et al. (2024) [51]	Used VR circuit training with feedback to boost skills and SRL among low performers.	Balance between guidance and autonomy remains underdefined for scalable use.
Wu et al. [52]	Created VR field simulations for landscape students to simulate design experience during remote learning.	Transition support and time for SRL adaptation in VR still under-researched.
Yang et al. [53]	Designed a VR CPR training system using feedback and self-paced practice.	More evidence needed on long-term retention and training frequency in self-paced VR.
Koçdar et al. [54]	Developed and validated a five-factor SRL scale for open/distance learning.	Needs application and validation within VR-specific settings.
Wang et al. (2020) [55]	Reviewed DL in VR content creation, enabling adaptation and personalization.	Gap in linking DL-based adaptation to measurable SRL strategies.
Villena-Taranilla et al. [4]	Meta-analyzed VR in K-6 education; found immersion effective especially in short sessions.	SRL potential in younger learners requires more focused exploration.

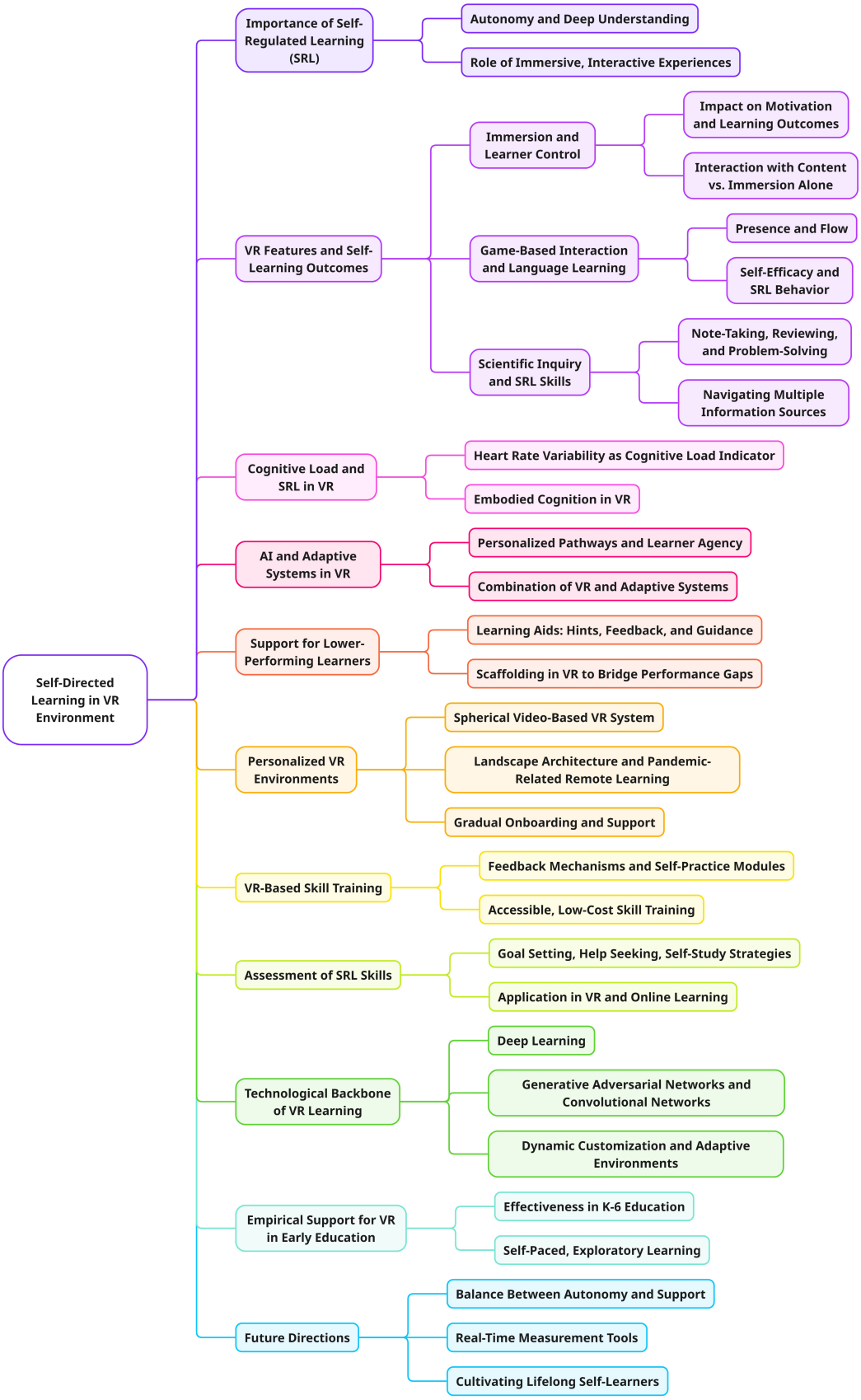


Figure 2.4: Thematic mind map summarizing self-directed learning in VR. Key branches include SRL mechanisms (e.g., autonomy, self-efficacy), VR features (e.g., immersion, adaptive systems), learner characteristics (e.g., performance level), and support strategies. The map visualizes interlinked research directions and challenges.

2.5 Mathematics Learning in VR Environment

Virtual reality (VR) has emerged as a promising medium for enhancing mathematics education by enabling interactive and spatially rich learning experiences. Its immersive nature allows for intuitive understanding of abstract concepts, particularly in topics involving three-dimensional geometry and multivariable calculus. A growing body of research has explored how VR can support students' mathematical reasoning, visualization skills, and motivation, though outcomes vary depending on system design, instructional strategy, and learner characteristics.

In the area of geometry education, Su et al. [1] developed a VR immersive learning system to teach core geometric topics such as triangular pyramid volume, cone volume calculation, and the center of gravity of a triangle. The system enabled students to explore and manipulate geometric structures in a virtual environment, fostering deeper conceptual understanding. Experimental results showed that students who used the VR system achieved better learning outcomes and reported higher confidence and motivation compared to those receiving traditional instruction. The researchers suggest that incorporating VR into math classrooms can provide game-like elements and task completion satisfaction, enhancing both engagement and comprehension.

Similarly, Simonetti et al. [2] designed a VR platform to support the teaching of analytical geometry and physics concepts in high school. Their system emphasized the benefits of immersive environments over conventional two-dimensional methods, citing improved spatial reasoning and reduced cognitive fragmentation. While the authors stressed that VR should complement rather than replace traditional instruction, their findings suggest that well-structured VR tools can significantly strengthen students' conceptual understanding and interest in mathematics.

VR has also shown potential in supporting the learning of advanced topics such as multivariable calculus. Cheong et al. [3] investigated a VR application aimed at improving students' ability to visualize partial derivatives and 3D surface behavior. Their study involved 119 engineering students and compared quiz-based learning with VR-enhanced instruction. While most learning outcomes remained similar across groups, students using the VR tool showed significantly better performance in identifying the sign of partial derivatives. This effect was especially pronounced among low-progress students, indicating that VR can provide additional support for learners who struggle with abstract mathematical concepts. The authors argue that such tools may contribute to developing future-ready smart classrooms by making complex calculus content more accessible.

However, research findings on the effectiveness of VR in mathematics education are not uniformly positive. Kang et al. [6] conducted a large-scale randomized controlled trial with 312 engineering students to evaluate a VR system designed for multivariable calculus instruction. Although students using the VR application reported positive learning experiences, they performed worse on certain test items compared to the control group. The authors attributed these outcomes to issues such as unstructured content and insufficient exposure time. They emphasize the need for careful instructional design, especially when using VR to teach complex subjects.

At a broader level, meta-analyses have investigated the general effectiveness of VR in improving learning outcomes. Villena-Taranilla et al. [4] synthesized findings from 21 experimental studies and found that VR-based instruction had a moderate positive effect on student performance, with an overall effect size of 0.64. Notably, immersive

VR systems outperformed semi- and non-immersive alternatives, particularly in primary education settings. The study also reported that shorter interventions (under two hours) were generally more effective than longer ones, suggesting that VR may be most beneficial when integrated into focused, modular learning experiences.

Hamilton et al. [5] conducted a systematic review focusing on immersive VR (I-VR) using head-mounted displays. Among the 29 publications analyzed, most reported learning gains associated with I-VR, although some studies found no significant difference compared to traditional methods, and only a few reported negative effects. The authors highlighted several common limitations, including short intervention durations, narrow focus on scientific topics, and limited attention to long-term knowledge retention. They concluded that more rigorous and standardized evaluation methods are necessary to fully understand VR's pedagogical value.

In terms of implementation, Marks and Thomas [56] evaluated the adoption of VR and augmented reality (VAR) in higher education over 2.5 years at the University of Sydney. Despite serving thousands of students, only a small percentage of engineering courses actively integrated VAR into their curriculum. Survey results indicated that most students perceived improved learning outcomes through the use of immersive tools. The authors concluded that although VR adoption remains in its early stages, investment in VR infrastructure can provide cost-effective and scalable opportunities for enhancing education through interactive 3D models and simulations.

Overall, existing research suggests that VR holds considerable promise as a supplement to mathematics instruction. Its strengths lie in improving spatial reasoning, supporting low-performing students, and enhancing learner motivation through immersive engagement. However, to fully realize its potential, educators and designers must ensure that VR systems are pedagogically grounded, well-structured, and appropriately timed within broader instructional frameworks. Future research should continue to explore how best to align VR technologies with curricular goals, learner needs, and assessment practices in mathematics education.

To complement the textual analysis, Figure 2.5 presents a thematic mind map of mathematics learning in VR. It visually organizes major branches such as geometry education, multivariable calculus, platform evaluation, and implementation challenges. The diagram highlights how research converges around key benefits (e.g., spatial reasoning, motivation), effectiveness (e.g., for low-performing students), and systemic concerns (e.g., curriculum integration and adoption).

Table 2.5: Summary and Identified Gaps in Mathematics Learning Research in VR

Study	**Summary**	**Gap**
Su et al. [1]	Developed an immersive VR system to teach geometric concepts such as volume and center of gravity. Students using the system showed improved performance, confidence, and motivation.	Although effective, the study focused on a limited set of geometric topics and lacked long-term assessment of knowledge retention and classroom scalability.

Study	Summary	Gap
Simonetti et al. [2]	Created a VR platform for analytical geometry and physics, enhancing spatial reasoning and reducing cognitive fragmentation.	The study primarily targeted high school learners and emphasized qualitative observations, with limited controlled quantitative validation across diverse curricula.
Cheong et al. [3]	Evaluated a VR tool for visualizing partial derivatives among engineering students. VR users outperformed in identifying signs of partial derivatives.	The tool's benefits were limited to specific problem types; the study did not evaluate broader conceptual understanding or real-world transferability.
Kang et al. [6]	Conducted a large-scale RCT on VR for multivariable calculus; while learners had positive experiences, some test scores were lower.	Findings raise concerns about poor instructional design and limited exposure time, suggesting the need for structured integration within curriculum.
Villena-Taranilla et al. [4]	Meta-analysis of 21 studies showed VR had moderate positive effects on learning, especially in short, immersive sessions.	Study emphasized primary education and short durations; less is known about effectiveness for advanced topics and longer-term use.
Hamilton et al. [5]	Systematic review of immersive VR in education using HMDs, reporting mixed but generally positive outcomes.	Many included studies had small sample sizes, lacked control groups, or did not measure long-term retention or real classroom applicability.
Marks and Thomas [56]	Evaluated real-world VAR adoption in higher education; students reported improved outcomes, though adoption was limited.	Despite positive perception, usage remained low due to infrastructure, training, and curricular integration barriers, highlighting the gap between feasibility and practice.

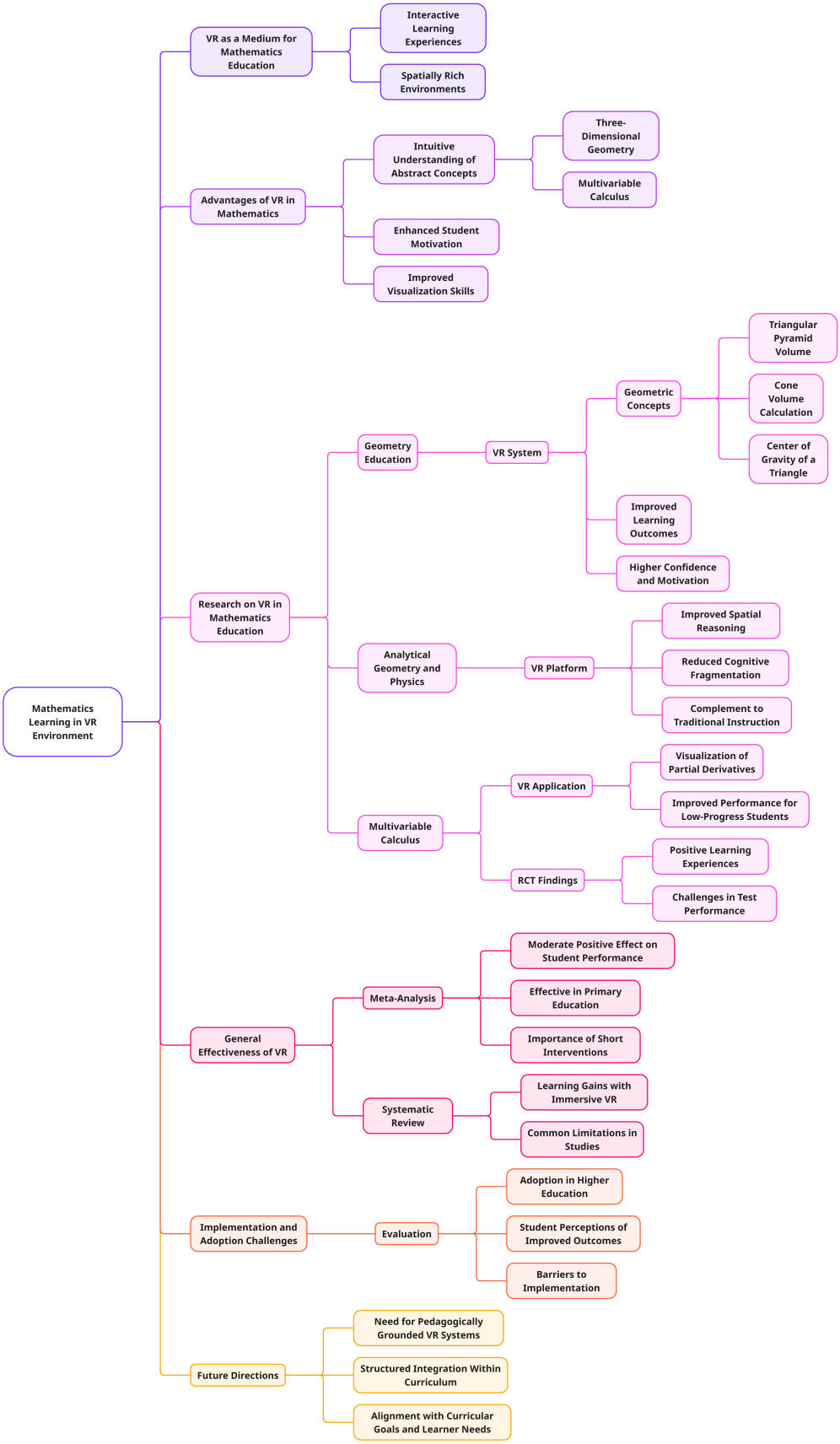


Figure 2.5: Thematic mind map summarizing research on mathematics learning in VR environments. The map outlines key areas including educational advantages (e.g., visualization, motivation), research findings in geometry and calculus education, evaluation outcomes, implementation challenges, and future directions.

2.6 Interface Design and Usability in Educational Technologies

Interface design plays a crucial role in shaping learners' cognitive processes, engagement, and overall educational outcomes. As digital learning environments become increasingly sophisticated, research has emphasized the importance of simplicity, usability, and clarity in interface design, especially when balancing interactivity and cognitive demand.

Gauthier et al. [57] investigated how different human-computer interaction (HCI) features in educational games influence young children's "stop-and-think" (S&T) behavior when solving counterintuitive science problems. Through a randomized eye-tracking study with 45 children aged 7 to 8, the researchers examined readiness mechanics, motion cues, color cues, and reward systems. The readiness mechanic and symbolic color cues both enhanced S&T duration and were associated with better in-game performance. In contrast, reward and penalty mechanisms appeared to distract students from reflective thinking. The study emphasizes the impact of subtle interface cues on intended cognitive engagement and highlights the need for thoughtful HCI design grounded in learning science.

In the domain of multimedia learning, Peng et al. [58] compared three interface design strategies: holistic positive emotional design, local positive emotional design (LPED), and neutral design. Their eye-tracking and survey-based study showed that LPED was most effective in promoting cognitive effort and focus, which ultimately led to better retention and transfer performance. These results suggest that incorporating emotional design at a local interface level, such as specific colors or visual accents, can meaningfully improve learners' attention and processing without overwhelming the user interface.

Interface simplicity is also linked to broader engagement and satisfaction outcomes. Pandita and Kiran [59] surveyed 400 students in higher education to examine how three dimensions of technology interface, namely cyber infrastructure, content quality, and technology-assisted facilities, affect student engagement (SE) and satisfaction (SS). Their structural model analysis revealed that interface design not only directly influences SE and SS, but also indirectly impacts satisfaction through engagement. The study underscores the importance of integrating user-friendly, high-quality digital systems with opportunities for active involvement and skill development.

Alzahrani and Roberts [60] examined how three types of multimedia interfaces, including menu-driven (MDI), graphical user (GUI), and zoomable user interfaces (ZUI), affect second language vocabulary acquisition. Their findings revealed that ZUI led to the highest immediate learning gains, but GUI proved more effective for long-term retention. This result suggests that interface design should consider both immediate usability and long-term knowledge consolidation, depending on learning goals.

Chatain et al. [61] focused on embodied interactions in mathematics education through tablet and VR technologies. They designed an interactive system to teach derivatives using varying degrees and forms of embodied interaction. Usability testing and learning outcome analysis revealed that simpler, more intuitive forms of embodiment improved understanding. This study supports the idea that while embodied learning can enhance concept grounding, overly complex interaction modalities may hinder learning unless carefully designed.

Wang et al. [62] explored whether task complexity in a VR learning environment

affects students' learning styles in construction education. Across three different tasks, their study with 253 undergraduates showed no significant influence of task complexity on learning style preference. This suggests that learners may adapt to VR interfaces consistently, regardless of how cognitively demanding the task is. These findings highlight the value of designing clear and consistent interfaces across various difficulty levels.

Finally, Scanlon [63] offered a reflective synthesis of 50 years of educational technology research. Through the "TEL Complex" framework, the study traces how educational interfaces have evolved through interdisciplinary influences such as machine learning, social learning, and personalization. The author argues that although technology has grown more advanced, simplicity and transparency in design remain central to achieving meaningful educational outcomes, particularly in distance and adult education.

Taken together, these studies underscore that effective interface design in educational technologies is not merely a technical concern but also a pedagogical one. Simplicity, emotional resonance, and usability must be balanced with functionality and learner needs. As educational platforms continue to integrate immersive, embodied, and data-driven technologies, maintaining clear, focused, and user-friendly interfaces will remain critical to fostering engagement and facilitating learning.

To visually summarize the thematic scope of this section, Figure 2.6 presents a mind map of research on interface design and usability in educational technologies. The diagram captures core themes such as HCI features, interface simplicity, emotional design, embodied interaction, and directions for future research, reflecting their interconnected roles in fostering student engagement and cognitive effectiveness.

Table 2.6: Summary and Identified Gaps in Interface Simplicity Research in Educational Technologies

Study	**Summary**	**Gap**
Gauthier et al. [57]	Studied interface features in educational games for young children solving counterintuitive science problems. Readiness mechanics and symbolic color cues enhanced stop-and-think behavior.	Needs deeper investigation on how interface design cues impact metacognitive strategies in broader age groups and content domains.
Peng et al. [58]	Compared emotional interface designs in multimedia learning. Local positive emotional design improved attention and retention.	More research needed on scalable implementation of emotional design in complex learning systems.
Pandita and Kiran [59]	Analyzed how technology interfaces affect student engagement and satisfaction in higher education through structural modeling.	Requires longitudinal studies to assess how interface-driven engagement affects learning outcomes over time.

Study	Summary	Gap
Alzahrani and Roberts [60]	Compared menu-driven, GUI, and ZUI interfaces for vocabulary learning. GUI aided long-term retention; ZUI led to immediate gains.	Further exploration needed on aligning interface types with specific learning phases and retention goals.
Chatain et al. [61]	Designed a math learning system using VR and tablet-based embodied interaction. Simpler interfaces improved usability and learning.	More work needed on balancing embodiment and simplicity for sustained learning in STEM contexts.
Wang et al. [62]	Studied how VR task complexity affects learning styles in construction education. Found consistent adaptation regardless of task difficulty.	Need to investigate how to personalize interface clarity without oversimplifying complex skill domains.
Scanlon [63]	Provided a reflective analysis of 50 years in technology-enhanced learning design, emphasizing simplicity and transparency.	Future research should operationalize simplicity across emerging AI-enhanced and immersive systems.

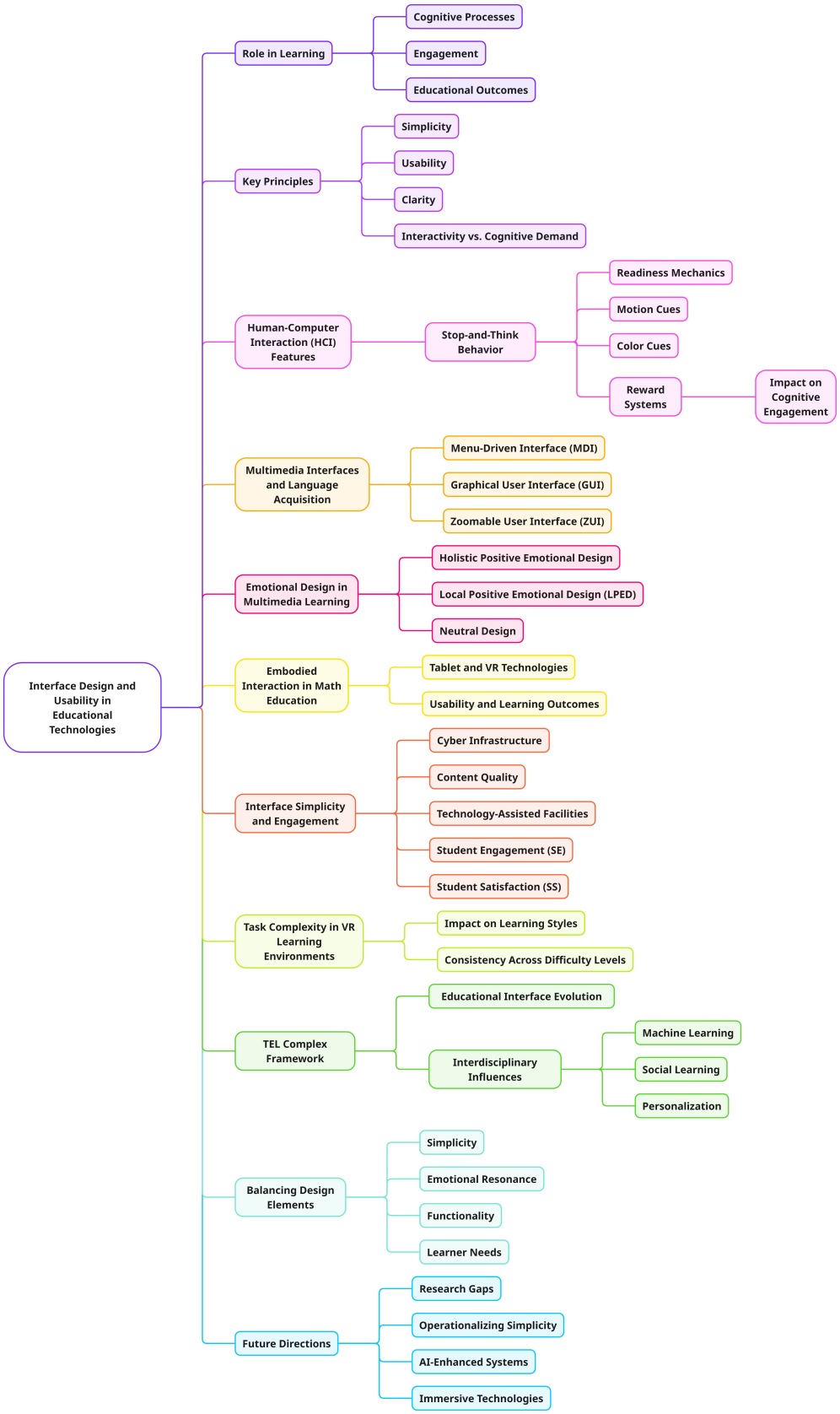


Figure 2.6: Thematic mind map of interface design and usability in educational technologies. Key branches include principles of simplicity and usability, cognitive and emotional design strategies, multimodal interaction approaches, and system-level considerations in educational platforms.

2.7 Instructor Presence and Cognitive Load in Educational Media

The impact of instructor presence and visual embodiment in instructional media has been widely studied, particularly in relation to cognitive load, motivation, and learning outcomes. Although numerous studies suggest that a visible instructor can enhance learner engagement and social presence, findings regarding its effects on cognitive processing and academic performance remain mixed.

Kizilcec et al. [64] conducted both observational and experimental studies to examine the role of instructor visibility in online video lectures. In their 8-week observational study involving 2,951 learners, students who chose to view the instructor's face reported a more engaging experience, but a notable 35% opted out to avoid distractions. A subsequent 10-week experiment with 12,468 learners compared constant versus strategic face presentation. Although strategic presentation increased social presence and perceived cognitive load, it did not lead to differences in attrition or learning outcomes. Notably, students with a verbal learning preference experienced lower cognitive load and lower dropout rates when the instructor's face was constantly present. These results support the image principle in multimedia learning but also highlight the trade-off between social presence and cognitive processing.

Beege et al. [65] conducted a meta-analysis of 35 studies with over 6,000 participants to examine the effect of visible instructors in instructional videos. They found that including an instructor led to small but statistically significant gains in retention, although transfer performance remained unaffected. Visible instructors enhanced social presence and motivation and reduced perceived extraneous cognitive load. However, they also decreased learners' visual attention to instructional content, suggesting a potential distraction effect. Moderator analyses indicated that prior knowledge, content domain, and screen size influenced these outcomes.

Polat [66] reviewed 41 empirical studies to explore how instructor presence influences learning in instructional videos. While students generally responded positively to instructor visibility, the results were not consistently linked to improved academic performance. Instructor presence, especially in formats with direct gaze and picture-in-picture layout, was associated with stronger attentional and social cues. The review concluded that further research is needed to clarify how instructor presence functions across different content areas and learner profiles.

Ng and Przybyłek [67] offered a more nuanced perspective by investigating vocabulary learning outcomes in two groups exposed to videos with and without visible instructors. Their findings revealed no significant differences in cognitive load, learning outcomes, or social presence between the two conditions. The authors proposed that individual differences, such as motivation and language aptitude, may have overridden the effects of instructor visibility. They also noted that the measures used may not have been sensitive enough to detect subtle cognitive effects in a between-subjects design.

Yu [68] explored the relationship between teacher presence, intrinsic cognitive load, and academic performance. Through a mixed-method study, the researcher found that videos featuring teacher presence significantly increased both intrinsic load and academic achievement. Intermittent teacher presence yielded greater benefits than constant presence, suggesting that a balanced presentation strategy may be more effective than a static format. The study calls for future research using multimodal and interdisciplinary approaches to further unpack these dynamics.

From a theoretical perspective, Castro-Alonso et al. [69] reviewed five instructional design strategies based on cognitive load theory, including multimedia integration, signaling, and segmentation. They proposed that the effectiveness of instructor guidance may vary depending on learner expertise. Specifically, novices benefit more from instructor-managed learning, while experienced learners may perform better when given control. This insight aligns with the expertise reversal effect and suggests that instructor presence should be adapted to learner proficiency levels.

Wenk et al. [70] examined cognitive and motivational responses to visualization technologies, comparing immersive virtual reality (IVR), augmented reality (AR), and 2D screens. Their results showed no significant difference in cognitive load across conditions, but IVR yielded higher motivation and usability ratings. IVR and AR also provided greater embodied learning experiences than 2D screens. These findings indicate that immersive environments can enhance the affective and usability dimensions of learning, which may interact with cognitive load in complex ways.

In a related study on online teaching presence, Wang and Stein [71] found that instructional design and facilitation increased reflection and cognitive engagement among students. Through a three-cycle course design with 132 participants, the researchers concluded that cognitive conflict, prompted by teacher presence, contributed positively to learning. Interestingly, direct instruction did not significantly alter students' responses to conflict, suggesting that structured interaction may be more effective than content delivery alone.

Finally, Ahuja et al. [72] investigated how pedagogical agents affect learners with and without dyscalculia. Their intelligent tutoring system used personalized paths and measured three types of cognitive load. The study revealed that well-designed agents did not increase intrinsic or extraneous cognitive load and instead enhanced germane cognitive load and retention. This demonstrates that instructional presence, when carefully implemented, can benefit diverse learner populations without causing cognitive overload.

Collectively, these studies illustrate that instructor presence and visual embodiment can shape learners' motivation, engagement, and cognitive load in different ways. While a visible instructor can improve social presence and retention under certain conditions, its impact on cognitive processing and academic outcomes is moderated by factors such as learner expertise, content type, and instructional strategy. Educational designers should therefore consider individual differences and task characteristics when integrating instructor presence into video-based and immersive learning environments.

To further illustrate the multidimensional impact of instructor presence on learning processes, Figure 2.7 presents a mind map that organizes major research themes, including effects on cognitive load, engagement, learning outcomes, moderating variables, and technology-mediated factors.

Table 2.7: Summary and Identified Gaps in Instructor Presence and Cognitive Load Research

Study	Summary	Gap
Kizilcec et al. [64]	Compared constant vs. strategic presentation of instructor faces in video lectures; strategic visibility increased perceived cognitive load but did not improve learning outcomes.	Further work needed to disentangle the benefits and costs of showing the instructor's face in different learning contexts.
Beege et al. [65]	Meta-analysis of 35 studies showed visible instructors slightly enhanced retention and motivation, but also decreased attention to instructional content.	More investigation needed into the balance between social presence and distraction, and how it varies with learner background and screen size.
Polat [66]	Systematic review of 41 studies on instructor visibility in videos found generally positive effects on engagement but inconsistent learning gains.	Needs further research to clarify mechanisms across domains and learner types, especially with eye contact and layout differences.
Ng and Przybyłek [67]	Compared vocabulary learning with and without visible instructors; found no significant differences in load or outcomes.	Suggests a need for more sensitive measures and consideration of individual factors like motivation and aptitude.
Yu [68]	Found that intermittent teacher presence improved academic outcomes and intrinsic cognitive load more than constant presence.	Calls for future work using multimodal and interdisciplinary approaches to explain these effects.
Castro-Alonso et al. [69]	Reviewed cognitive load strategies and proposed guidance should vary by learner expertise.	Systematic testing of instructor- vs. learner-controlled designs needed across domains.
Wenk et al. [70]	Compared IVR, AR, and 2D screens; found higher motivation and usability in immersive modes, but no cognitive load differences.	Needs further work on how embodied experience and media form interact with load and affect.
Wang and Stein [71]	Found teaching presence and instructional design boosted engagement and reflection in online settings.	Suggests structured interaction may be more effective than content delivery alone; impact of direct instruction remains unclear.
Ahuja et al. [72]	Studied pedagogical agents with students with and without dyscalculia; enhanced germane cognitive load and retention.	Future work should explore how tailored agents can benefit neurodiverse learners without raising load.

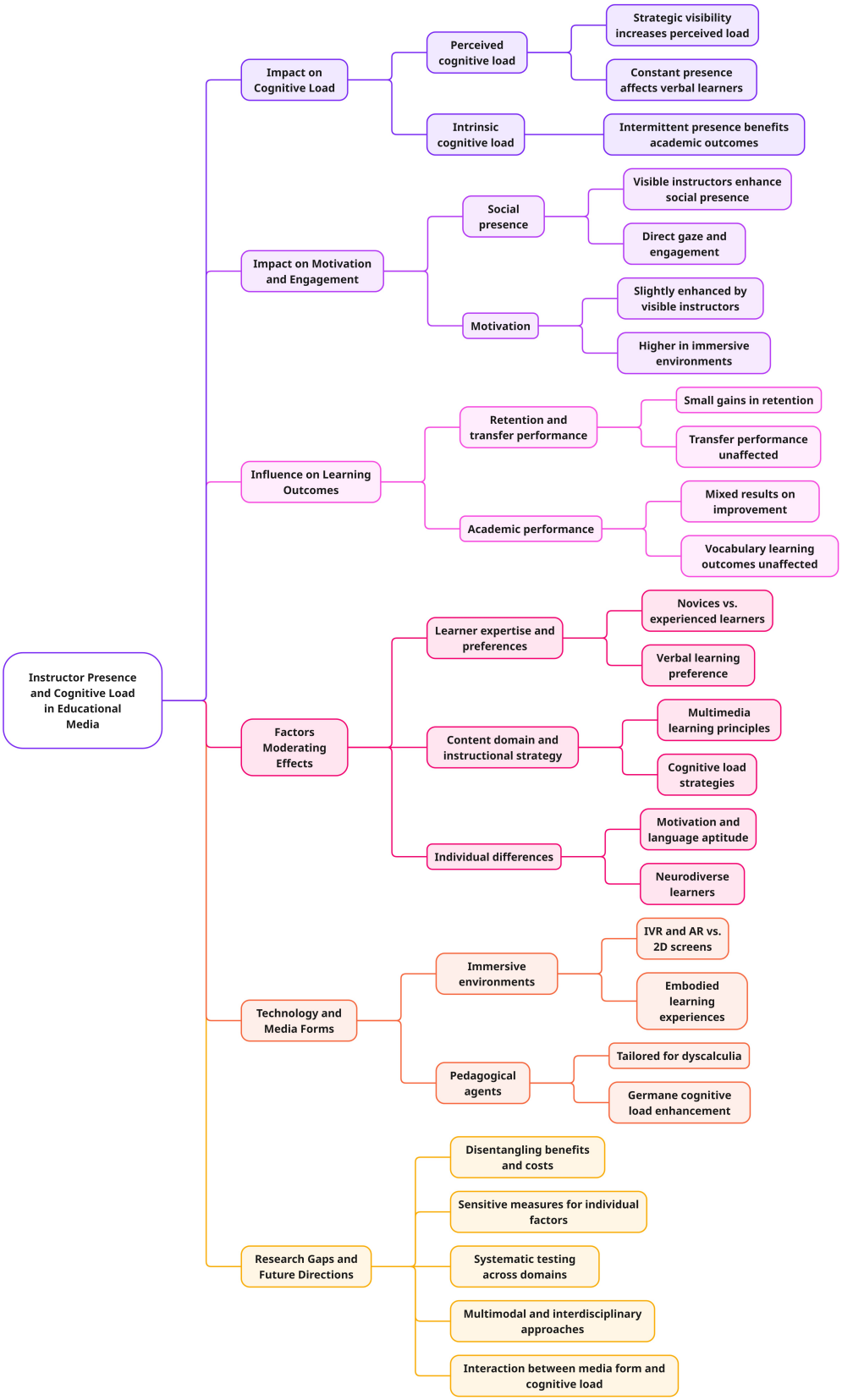


Figure 2.7: Mind map of research themes related to instructor presence and cognitive load in educational media. Key domains include cognitive and motivational effects, individual and contextual moderators, and the role of immersive environments and pedagogical agents.

Chapter 3

System

This chapter presents the design and technical composition of the VR-based self-learning system developed for this study. It begins by describing the original vision and how the design evolved over time, followed by a detailed explanation of the implemented prototype, including its main modules, development environment, and platform-specific adaptations. The goal is to provide readers with a comprehensive understanding of how the system was conceptualized, built, and structured to support immersive AI-assisted learning in a virtual environment.

3.1 Original Vision and Design Evolution

The system presented in this thesis was originally envisioned as a 3D adventure-puzzle game designed to teach higher-level mathematics, specifically calculus, through immersive gameplay and narrative. Rather than teaching basic arithmetic or geometry, which are commonly found in educational games, the core idea was to make abstract and often intimidating concepts in advanced mathematics more engaging and accessible through exploration-based mechanics such as manipulating spatial objects, solving navigation puzzles, and observing interactive environments to derive abstract rules.

This direction was conceived independently, and only later did I discover the project *Variant: Limits* [73], as shown in Fig.3.1, a 3D puzzle game that closely aligns with my original goals. Their implementation confirmed the feasibility of combining spatial puzzles with the calculus curriculum, serving as a useful reference in articulating and refining my own design goals.

The early design centered around a first-person VR environment where players would explore a stylized world to uncover calculus principles embedded in the environment. Tasks included manipulating geometric shapes to visualize limits, derivatives, or integrals, solving spatial puzzles to unlock new mathematical insights, and interacting with a personalized AI character for hints and explanations. To support modularity and future extensibility, the system architecture was originally planned to allow for incremental addition of new modules or storylines.

A crucial element of this vision was a highly interactive AI avatar who would act as a guide, tutor, and companion throughout the gameplay. The inspiration for this component came from an online modification to *The Elder Scrolls V: Skyrim*, where a community developer integrated ChatGPT into a game NPC, allowing the character to converse freely with the player [74]. As shown in Fig.3.2, this example demonstrated

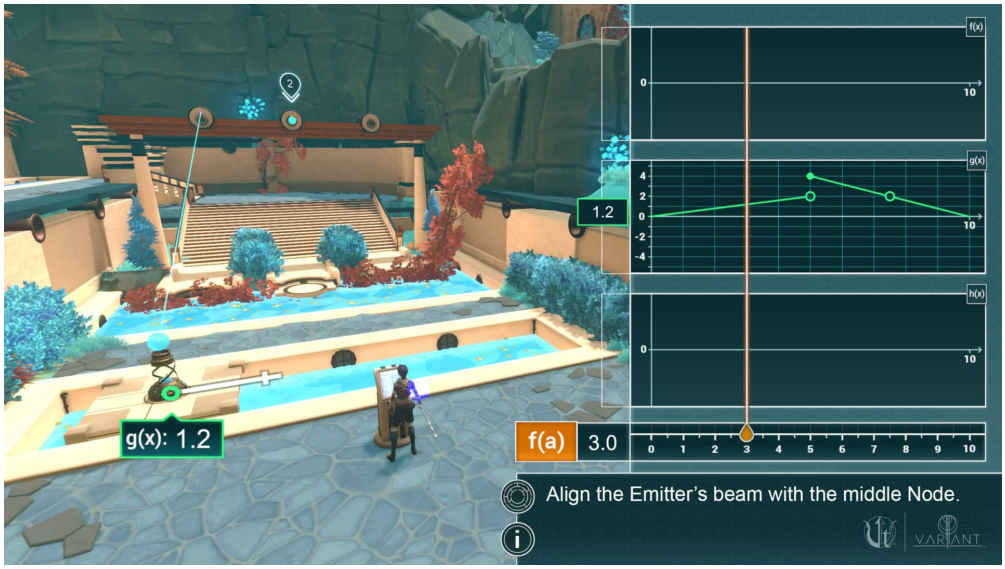


Figure 3.1: A screenshot from the educational game *Variant: Limits*, showing a puzzle scene where the player uses graphical function data to align an emitter with a target node. On the right side of the interface, three function plots are displayed—$f(x)$, $g(x)$, and $h(x)$—with a vertical line marking the input $a = 3.0$. The current value of $g(3.0) = 1.2$ is indicated both on the graph and in the 3D scene, where the player must match the emitter beam to the corresponding node based on this output. This visual integration of mathematical function evaluation with spatial interaction exemplifies how calculus concepts are applied in a game-based learning environment.

the powerful potential of merging generative AI with narrative environments, which inspired the concept of a math-focused VR assistant with its own personality.

To implement this avatar, I initially explored using a diffusion-based model to generate facial expressions dynamically based on emotional tone and context. However, the complexity of real-time expression synthesis exceeded the technical scope of this thesis. As an alternative, I began testing Inworld AI [75], which allowed for the creation of emotionally expressive and customizable virtual agents. The tool showed promising results during early experiments, but the platform's transition to a commercial license rendered it inaccessible during the development phase. Consequently, I migrated to OpenAI's Assistant API, which provided greater control and affordability.

I later discovered Convai [76], a promising alternative to Inworld AI that supports real-time AI avatars with voice and emotional nuance. It is also free to use within a generous limit, with usage-based charges applied only beyond that. However, by the time of this discovery, the main system was already near completion. Although I briefly experimented with Convai in a separate prototype, integrating it into the existing system would have required significant restructuring and was therefore not pursued. Nevertheless, Convai remains a viable option for future iterations of this project, and its integration could be revisited to enhance avatar expressiveness and interactivity. These developments marked key decision points in shaping the system's AI communication module. A visual example of both Inworld AI and Convai in action is shown in Fig. 3.3, demonstrating how real-time AI avatars can be integrated into interactive applications.

In addition to gameplay and interaction, the original system was also intended to feature real-time 3D visualizations of calculus concepts. Several existing tools, as



Figure 3.2: An in-game scene from a modded version of *The Elder Scrolls V: Skyrim*, where a non-playable character (NPC) named Herika, enhanced with ChatGPT integration, engages in dynamic, voice-acted conversation with the player. This example illustrates the potential of combining generative AI with immersive environments to create lifelike, responsive characters. It served as a key inspiration for the development of a personalized AI companion in the envisioned VR-based calculus learning system.



Figure 3.3: Visual comparison between two advanced real-time AI avatar systems. The left side shows a character created with Inworld AI, which supports context-aware emotional expression, memory retention, goal-driven behavior, and narrative integration. It was widely demonstrated at GDC 2024 in collaboration with companies such as Ubisoft and Xbox. The right side presents a virtual museum guide powered by Convai, a lightweight alternative that enables real-time voice interaction, expressive gestures, and multi-turn conversations. Both systems allow developers to create lifelike AI-driven characters, with Inworld offering more sophisticated personality modeling and Convai focusing on accessibility and performance efficiency.

shown in Fig.3.4, such as CalcFlow [77], CalcPlot3D [78] and CalcVR [79], offered technical references for how abstract mathematical functions could be spatialized and manipulated in immersive environments. These tools helped shape early prototyping ideas, such as rotating solids to understand volume or dragging points on a graph to intuit derivatives. However, developing custom educational content proved challenging: although I had access to textbooks and online materials, I lacked formal training in math pedagogy, and designing a structured calculus curriculum from scratch would have diverted focus from system interaction research.

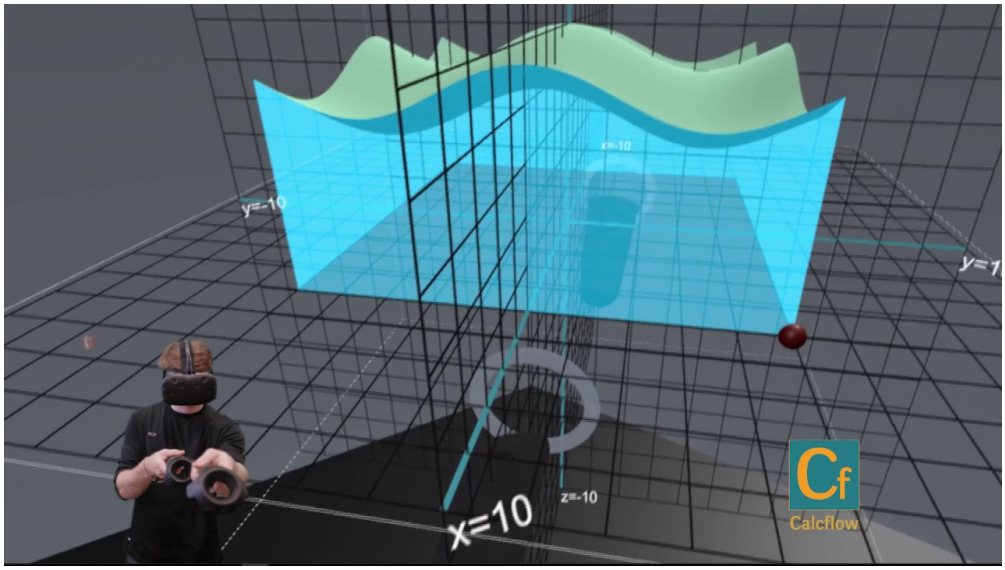


Figure 3.4: An example of real-time mathematical visualization in VR using the tool *CalcFlow*. The system allows users to interact with multivariable functions in 3D space by manipulating surfaces, planes, and coordinates. In this scene, a user wearing a VR headset explores a surface plot intersected by a transparent plane, aiding the conceptual understanding of partial derivatives or volume integration. Tools like *CalcFlow* served as technical inspiration for the early design of our system, particularly for embedding calculus visualizations in an immersive learning environment.

To maintain coherence and ensure content quality, I ultimately decided to adopt video materials from Khan Academy as the instructional core. This pivot allowed me to focus development on the surrounding VR interaction framework, which includes avatar communication, speech input and output, and quiz-based evaluation. This ensured reliable mathematical instruction without the need to design a full curriculum from scratch.

In the early stage of system evaluation, I employed Cognitive3D [80] to collect user interaction data during pilot testing. This platform provided useful tools for tracking gaze, session flow, and spatial analytics in VR, which made it appealing for early experiments. As shown in Fig. 3.5, the system allowed me to visualize user movements in the 3D scene and review interaction timelines using Scene Explorer. It also offered a project dashboard that summarized metrics such as comfort, presence, and performance scores, which helped identify possible friction points in the learning experience.

However, as the virtual environment became more complex and feature-rich, the web-based dashboard of Cognitive3D began to suffer from noticeable performance issues, including lag and frequent freezing during data visualization. In addition, some

advanced functionalities were difficult to configure without extensive documentation or high-end hardware. These limitations made the tool impractical for continuous use during later stages of development. As a result, I decided to remove Cognitive3D from the final version of the system and instead implemented a lightweight, custom logging method. This alternative approach offered more control, reduced system load, and simplified data retrieval for analysis, while still capturing essential user actions such as speech interactions and quiz responses.

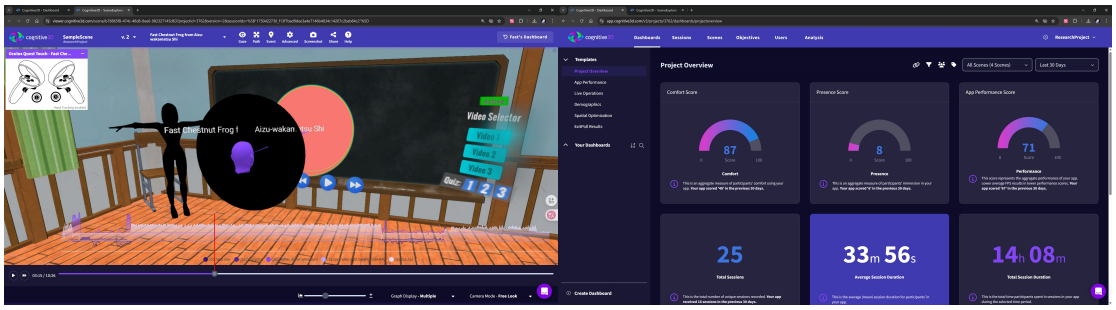


Figure 3.5: Early pilot test using Cognitive3D. The left side shows the Scene Explorer, which replays user movement and interaction paths in the VR environment along a timeline. The right side displays the analytics dashboard, with scores for comfort, presence, and app performance. This interface was useful for identifying bottlenecks and guiding early-stage system adjustments.

In sum, the final system reflects a pragmatic compromise between ideal vision and development feasibility. While many original features were not implemented due to time, technical, or educational constraints, they continue to inspire directions for future research and expansion. The design trajectory also illustrates how interdisciplinary challenges, involving AI, animation, curriculum design, and HCI, shaped the project into its current form.

3.2 Prototype Overview

We developed a modular virtual reality (VR) learning system to support interactive, self-paced mathematics education. The system allows learners to enter a virtual classroom where they can watch short instructional videos, ask questions to an AI assistant using speech, and complete follow-up quizzes. The prototype was developed using Unity and tested on a Lenovo Legion laptop before being deployed to a Meta Quest 3 headset, as shown in Figure 3.6.

Unity was selected for its flexibility and rich VR development ecosystem, while the Meta Quest 3 provided a comfortable and accessible platform for immersive testing. The system delivers short instructional videos, followed by quizzes and an AI assistant that supports spoken user input. Speech is transcribed using Whisper and processed by GPT-4o-mini to generate text-based responses. When the embodied avatar is active, the responses are also converted to speech using Amazon Polly. This setup enables flexible interaction modes, allowing users to choose between text-only or multimodal communication depending on their preference.



Figure 3.6: Development and deployment setup. On the left is a Lenovo Legion Y9000X laptop running the Unity editor with the virtual classroom scene. On the right is the Meta Quest 3 headset with two white controllers, used for testing the VR learning experience.

As shown in Figure 3.7, the VR classroom consists of two main functions: (1) **Video-based learning** (Figure 3.8), and (2) **AI-based interaction** (Figure 3.9). On the right side of the classroom, users can select and play one of three short videos explaining calculus concepts. After watching, they select a quiz to test their understanding. On the left side of the classroom, users can interact with an AI assistant using either a text-based panel or an embodied 3D avatar. All activities are recorded locally for later analysis.

All interactions are logged locally, enabling researchers to analyze user behavior without depending on external analytics platforms. The complete source code for the Unity-based prototype has been made publicly available on GitHub [81], allowing educators and researchers to further adapt or extend the system for AI-driven VR education.

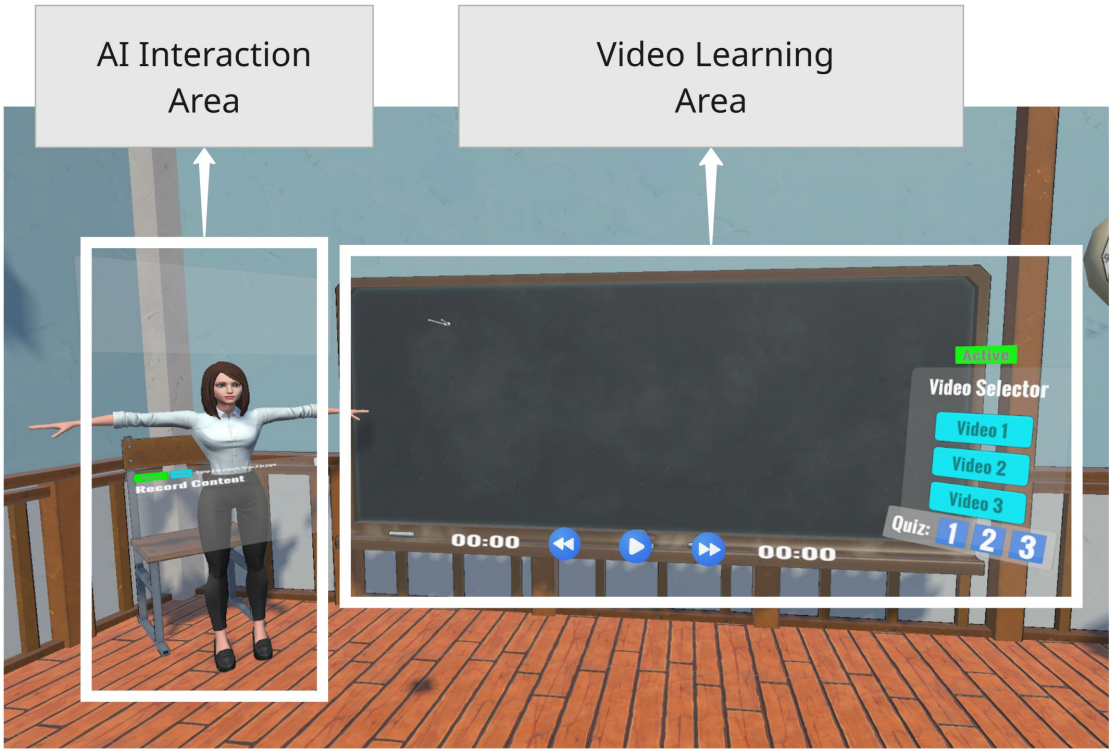


Figure 3.7: Overview of the VR learning system from the user's perspective. On the left, the AI interaction area includes a lower input panel for speech recognition, an upper response panel, and an embodied avatar. On the right, there is a video selector panel, a quiz selector panel, and a green "Active" button. When the "Active" button is toggled to "Hide," the avatar disappears. At the center, a video player is displayed with playback controls and time indicators. A quiz panel appears after clicking the quiz button, although it is not shown in this image.

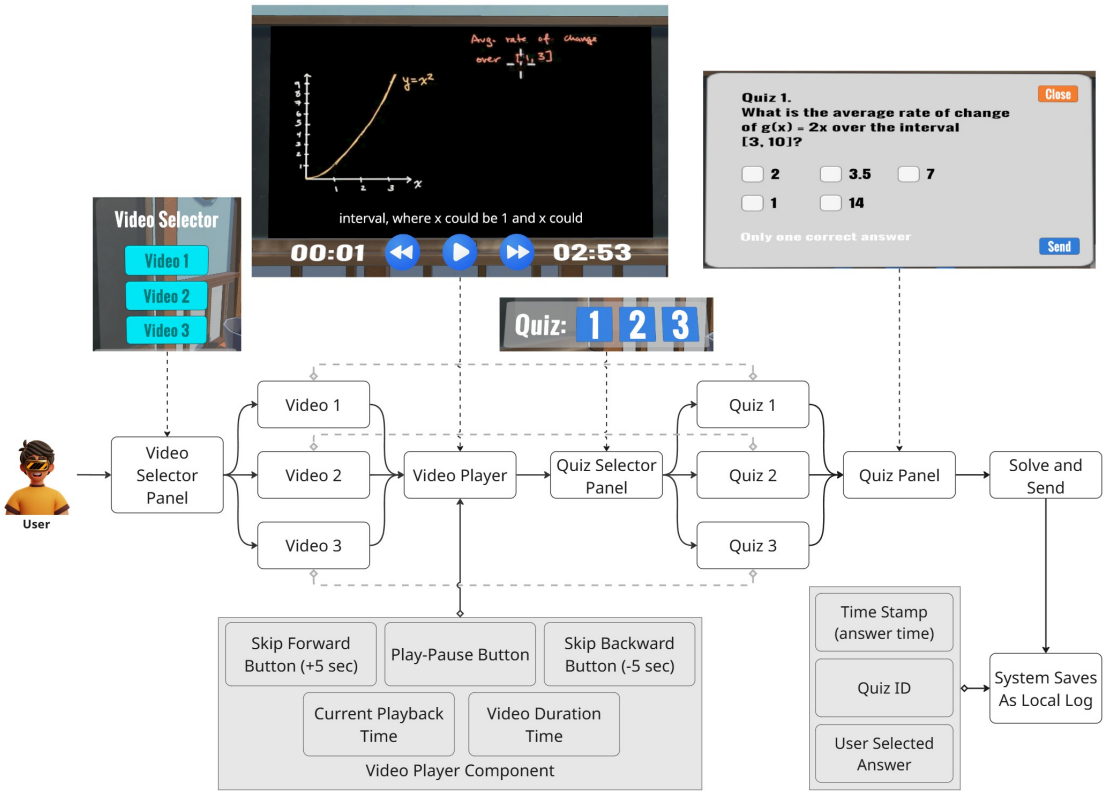


Figure 3.8: Video-based learning flow in VR. The user begins by selecting one of three calculus videos (Video 1, 2, or 3) through the **Video Selector Panel** on the left. The selected video is then loaded into the **Video Player**, which occupies the center of the classroom. Playback is controlled using three buttons located below the video: a play/-pause button, a skip forward button (+5 seconds), and a skip backward button (-5 seconds). The player also displays both the current playback time and the total video duration, as shown in the lower gray box labeled **Video Player Component**. After watching a video, the user clicks one of the numbered buttons (1, 2, or 3) on the **Quiz Selector Panel** to open the corresponding quiz. The selected quiz is shown in the **Quiz Panel**, where the user selects one answer from multiple options. The panel includes a "Send" button to submit the response and a "Close" button to exit the panel. Once the user clicks "Send," the system stores the quiz answer along with its timestamp and quiz ID in a local log file, as represented by the gray box labeled **System Saves As Local Log**.

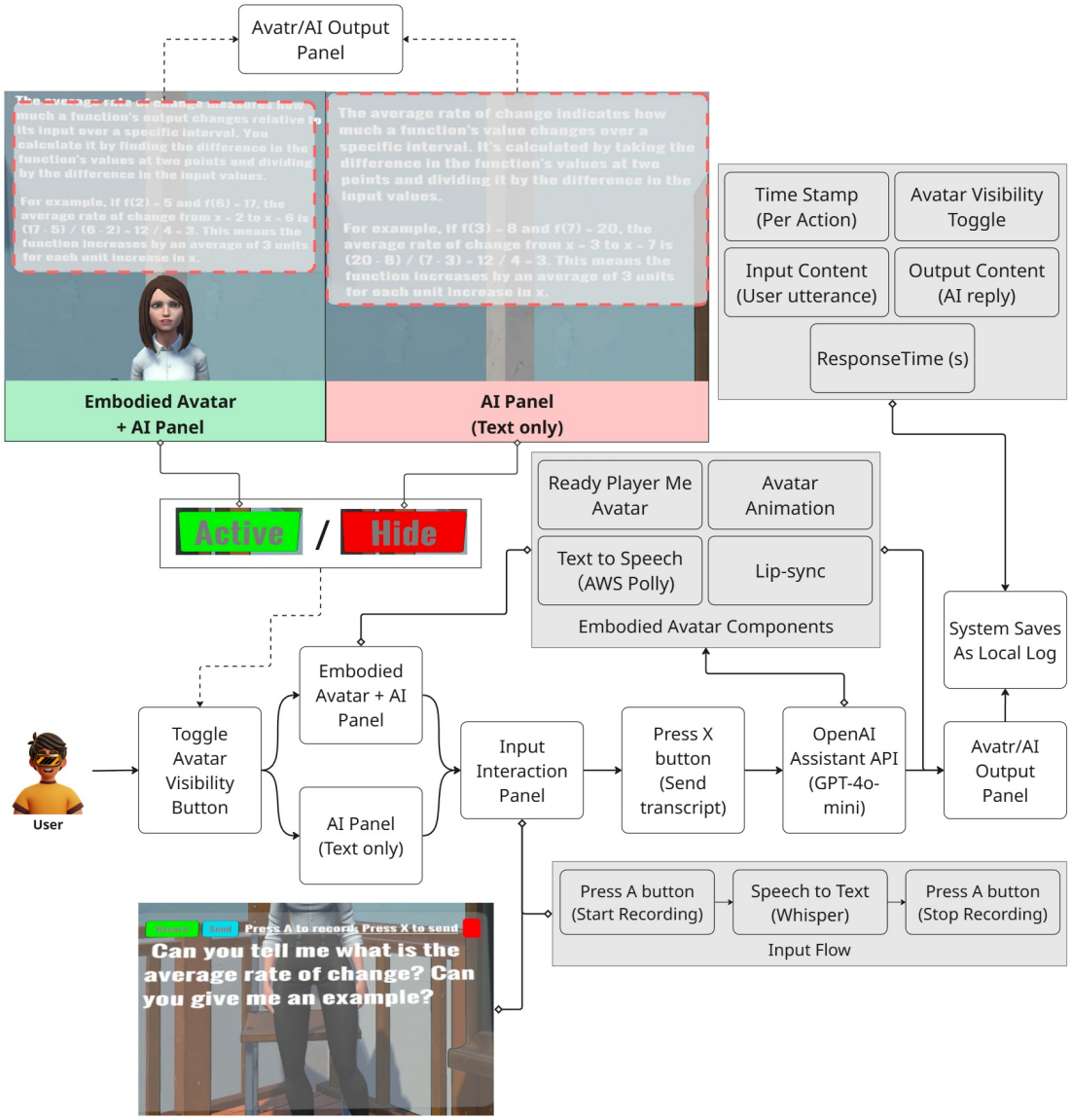


Figure 3.9: AI-based interaction flow in VR. The user interacts with the AI assistant through voice input using either the **Embodied Avatar + AI Panel** mode or the **Text-only AI Panel**. This can be toggled using the green/red button labeled ”Active” / ”Hide.” In either mode, the user initiates speech input by pressing the ”A” button to start recording and pressing it again to stop. The recorded audio is transcribed into text using the **Speech-to-Text (Whisper)** component. The transcribed input is then submitted to the AI assistant by pressing the ”X” button. This sends the query to the **OpenAI Assistant API (GPT-4o-mini)**, which generates a response. The output is displayed in the text panel and, if the avatar is active, converted to speech using **Amazon Polly**, animated via **Lip-sync**, and rendered by the **Ready Player Me Avatar** with predefined **Avatar Animations**. All interactions are logged locally by the system. The **System Saves As Local Log** module stores the avatar visibility state, user utterance, AI response, and response time in seconds, as shown in the upper right gray box. This ensures a complete record of each interaction step.

3.3 Implementation

In this chapter, we describe the architecture and implementation of the VR learning system. The system consists of four main functional modules: the **Video Learning Module**, **AI Assistant Module**, **Embodied Avatar Module**, and **User Interaction Logging Module**. Each module plays a dedicated role in supporting user learning and interaction within the VR classroom.

Figure 3.10 illustrates the overall architecture of the system, with each module shown in a distinct background color. Within each module, **dark gray** boxes represent concrete functional components such as buttons, APIs, or panels, while **light gray** boxes group related components to indicate layout or hierarchy.

The **Video Learning Module** (purple) includes interfaces for selecting and playing instructional videos based on Khan Academy content. Users can select from three videos, control playback, and respond to corresponding quizzes.

The **AI Assistant Module** (yellow) enables voice-based interaction. Speech input is recorded through the Input Panel, transcribed using Whisper, and processed by a GPT-4o-mini assistant. The assistant's response appears as text or is spoken aloud depending on whether the avatar mode is active. A backend Python server (via Flask and the OpenAI SDK) handles all dialogue communication.

The **Embodied Avatar Module** (pink) visualizes the AI assistant as a human-like teacher. The avatar is built with Ready Player Me and animated using Mixamo. AWS Polly handles speech synthesis, while Oculus LipSync synchronizes mouth movements with audio playback.

The **User Interaction Logging Module** (blue) records detailed interaction logs for both quizzes and AI dialogue. These include timestamps, user answers, response times, and input/output content, which can be used for later analysis of learner behavior.

3.3.1 Development Environment and Platform

The system was developed using Unity 2022.3.62f1 (URP) and deployed on Meta Quest 3 (128 GB) via the Unity OpenXR plugin. All development and testing were conducted on a PC equipped with an Intel Core i5-12500H CPU, 16 GB RAM, NVIDIA GeForce RTX 3050Ti GPU, and a 512 GB SSD, with approximately 7 GB used for this project. A summary of the development and runtime environment is provided in Table 3.1.

To enable in-editor debugging and reduce iteration time, the Meta Quest Link [82] was used to connect the headset to the development PC. A high-performance Thunderbolt 5 (USB-C) cable was used to ensure reliable data transmission, high bandwidth (up to 80–120 Gbps), and smooth real-time rendering. In practice, lower-grade USB-C cables caused rendering instability and noticeable lag, making cable quality a critical factor during VR development.

Once connected via Quest Link, the headset mirrored the Unity Editor's Game view, allowing direct testing and fine-tuning without the need to build an Android APK. To establish the link, the user must first put on the headset and select the Meta Quest Link option from the system menu, which switches the device into PC VR mode. In this state, running the Unity project in Play mode causes the scene to render directly inside the headset. Figure 3.11 shows the headset successfully connected and recognized in the development environment. The system was intentionally built and run as a stan-

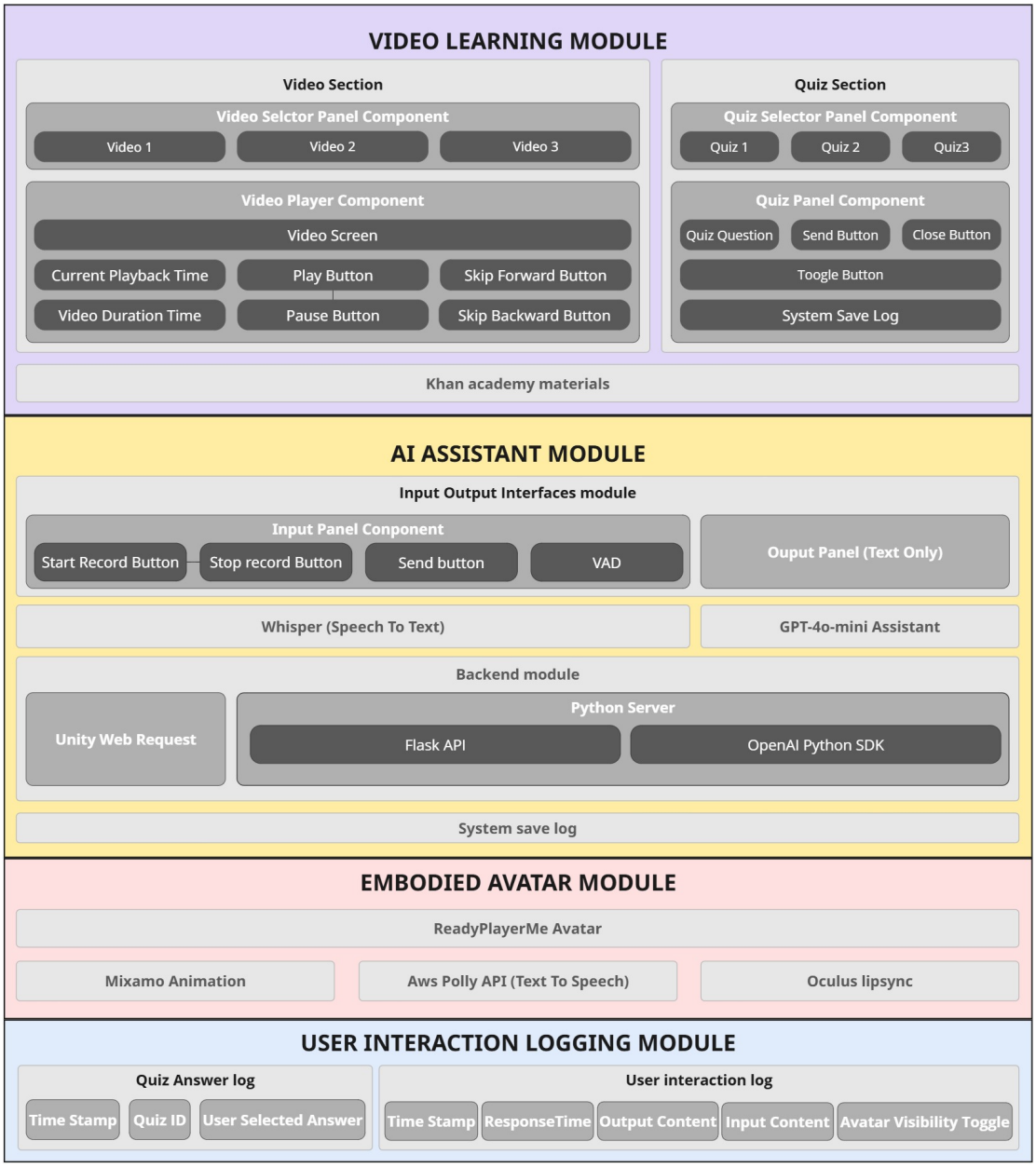


Figure 3.10: System architecture of the VR-based AI-supported learning environment, showing four color-coded modules and their internal components. Functional elements are shown in dark gray; layout groupings are in light gray. See main text for module descriptions.

dalone PC version because the Android build pipeline does not support local inference using Whisper models, which were essential for the AI assistant's speech recognition capability.

Prior to physically connecting the Quest 3, early-stage development was conducted using Unity's OpenXR Simulation mode. This simulation environment emulated headset movement and controller interactions, allowing rapid prototyping even without the physical device. This workflow proved especially helpful when working remotely or when connecting the headset was inconvenient, significantly improving development flexibility and iteration speed.

Table 3.1: Development and Runtime Environment Specifications

Component	**Specification**
CPU	Intel Core i5-12500H
RAM	16 GB
GPU	NVIDIA GeForce RTX 3050Ti
SSD	512 GB SSD (7 GB used for this project)
Operating System	Windows 11
Unity Version	Unity 2022.3.62f1 (URP)
Python Version	Python 3.13 (via PyCharm)
Headset Used	Meta Quest 3 (128 GB)
VR Runtime	Unity OpenXR Plugin + Meta Quest Link
Backend Framework	Flask (local server)

3.3.2 Video Learning Module

Three instructional videos were selected from Khan Academy's Differential Calculus series (Unit 2: Derivatives) [83]. Each video is approximately three minutes long and includes English subtitles. To ensure uniformity in length, all videos were trimmed using Clipchamp [84].

The selected videos are:

- **Video 1:** Secant lines and average rate of change [85].
- **Video 2:** Derivative as slope of a curve [86].
- **Video 3:** The derivative and tangent line equations [87].

These videos were integrated into the system using Unity's `Video Player` component [88]. Playback controls including play, pause, fast-forward (+5s), rewind (-5s), and duration tracking are implemented via a custom C# script (`VideoController.cs`). This script listens for user input from the VR controller and updates playback state accordingly. For example, pressing the Y button skips ahead 5 seconds, while pressing the B button rewinds.

Figure 3.12 shows the in-game video player interface, including the play/pause button, time display, and video selector panel.

Upon completing a video, the user can manually open a quiz by selecting one of the numbered buttons on the Quiz Selector panel. Each quiz includes five toggle options (radio buttons), from which the user selects one correct answer. The panel uses Unity's

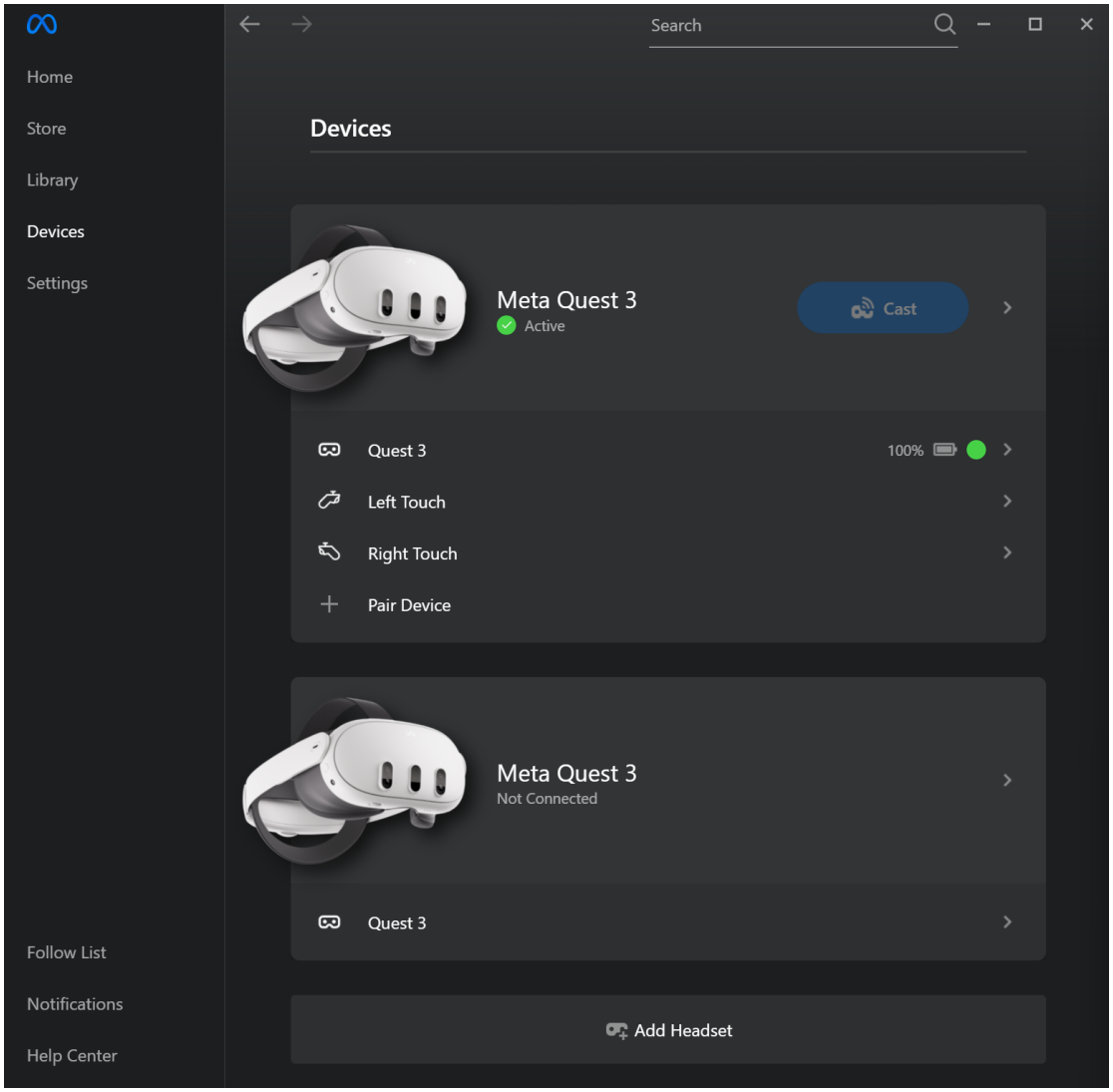


Figure 3.11: Meta Quest 3 successfully connected via Meta Quest Link. The development PC detects the headset and its controllers through a Thunderbolt 5 (USB-C) cable, enabling in-editor testing without Android builds. This setup allows real-time Unity execution and debugging directly within the PC environment.

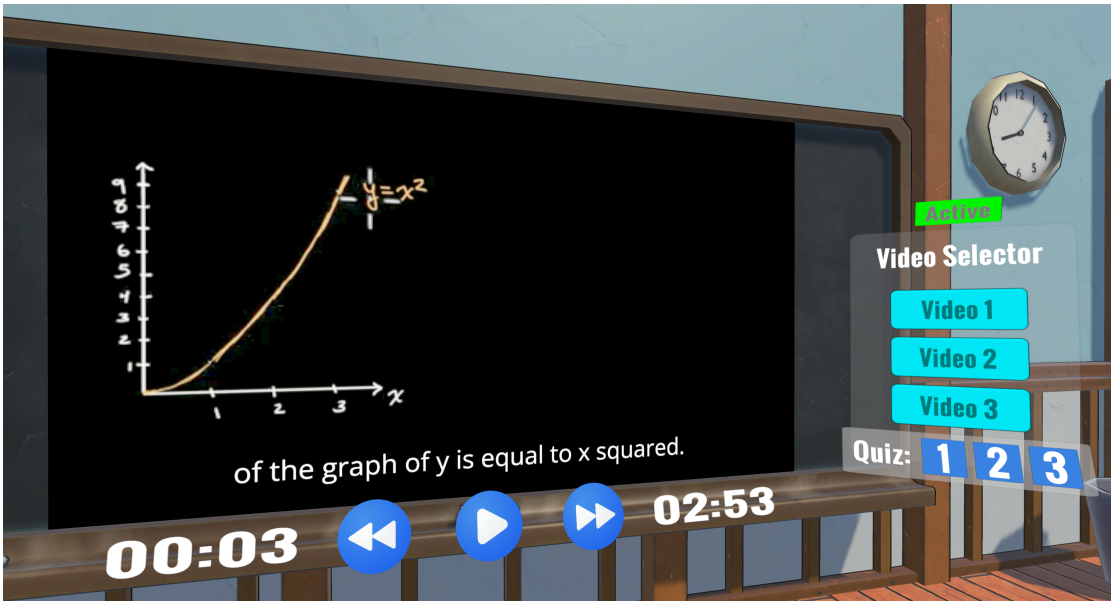


Figure 3.12: Unity-based video playback interface in the VR classroom. The selected calculus video appears at the center, with playback controls located below and a video selector panel on the right.

`ToggleGroup` system to ensure mutual exclusivity. A "Send" button records the response, and a "Close" button exits the panel. Upon submission, the response is saved locally in CSV format. The CSV entry includes the question string, timestamp, and selected answer.

Questions are adapted from corresponding Khan Academy exercises:

- **Quiz 1**: Secants and average rate of change [89].
- **Quiz 2**: Graphs of functions and their derivatives [90].
- **Quiz 3**: Derivative at a point as slope of tangent line [91].

Figure 3.13 shows an example of the quiz panel displayed after a video. The question appears at the top, followed by five answer toggles, and buttons to submit or close the panel.

3.3.3 AI Assistant Module

The AI interaction system supports two visual modes: a text-based AI panel and an embodied avatar interface. Users interact through a speech-based dialogue system implemented in Unity and powered by a customized assistant built using the OpenAI Assistants API [92]. This system integrates three major components: Whisper for speech-to-text transcription, the GPT-4o-mini model for response generation, and Amazon Polly for speech synthesis, which is only activated when the avatar mode is enabled.

Assistant Design and Training. The assistant was configured and tested using OpenAI's Assistants API Playground [92]. It was trained using selected video transcripts

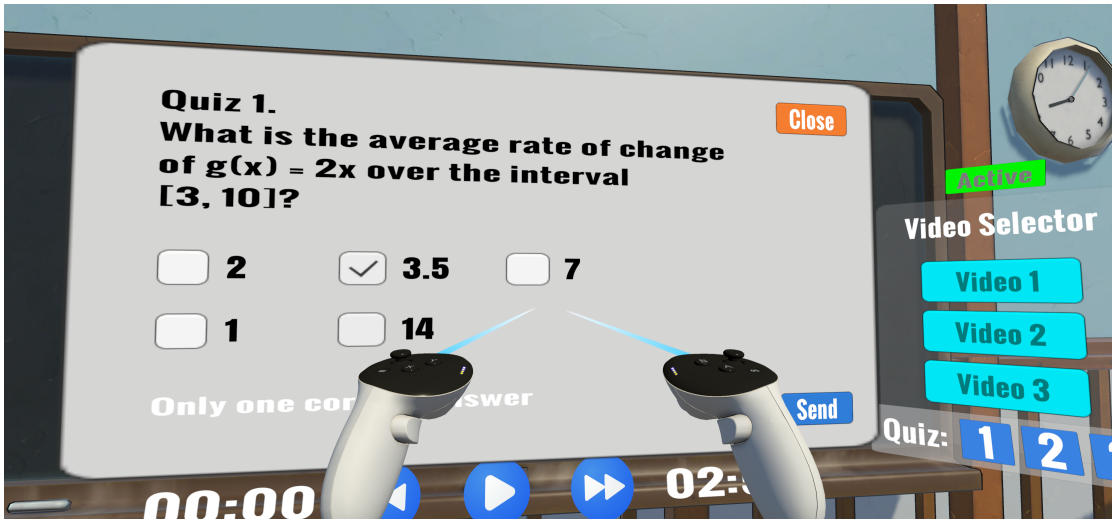


Figure 3.13: In-game view of the quiz interface. After a video finishes, the user can click one of the numbered buttons (1, 2, or 3) on the Quiz Selector panel to display the corresponding quiz panel. The quiz includes a multiple-choice question with five toggle options, rendered using Unity's ToggleGroup. The user selects one answer using the VR controller and clicks the "Send" button to submit. Submission automatically hides the panel and saves the quiz entry locally in CSV format. Alternatively, the user can close the panel by pressing the "Close" button without submitting. This figure also shows the surrounding interface, including the video selector and status indicator.

and quiz contents from Khan Academy's Unit 2 on Derivatives [83], including step-by-step explanations and correct answers. These materials were uploaded as system instructions and file attachments in JSON format. The assistant was designed to answer questions about calculus concepts, clarify content from the instructional videos, and offer hints related to quiz topics.

Runtime and Integration. The assistant runs on a locally launched backend developed in Python using Flask [93]. The backend is executed in PyCharm with Python 3.13 and listens at `http://127.0.0.1:5000`. Unity communicates with the backend via HTTP POST requests using the `UnityWebRequest` API [94], which relays queries to the OpenAI API through the official Python SDK [95]. Although the backend runs locally, it requires cloud-based OpenAI and Amazon Polly services for processing and audio synthesis.

Figure 3.14 shows the Flask server running inside PyCharm, handling incoming requests from Unity and dispatching them to OpenAI's API. To ensure proper connectivity between Unity and OpenAI services, this Python backend must be actively running in PyCharm before launching the VR application. If the server is not started, the avatar module in Unity will fail to retrieve responses and display an error. The full source code for the Python backend is publicly available on GitHub project repository [81].

Speech Input and Output. Users initiate voice input by pressing the A button on the Meta Quest controller. Unity records audio using the `Microphone` API [96], and transcribes it using the `whisper.unity` plugin [97], which runs the `base.en`

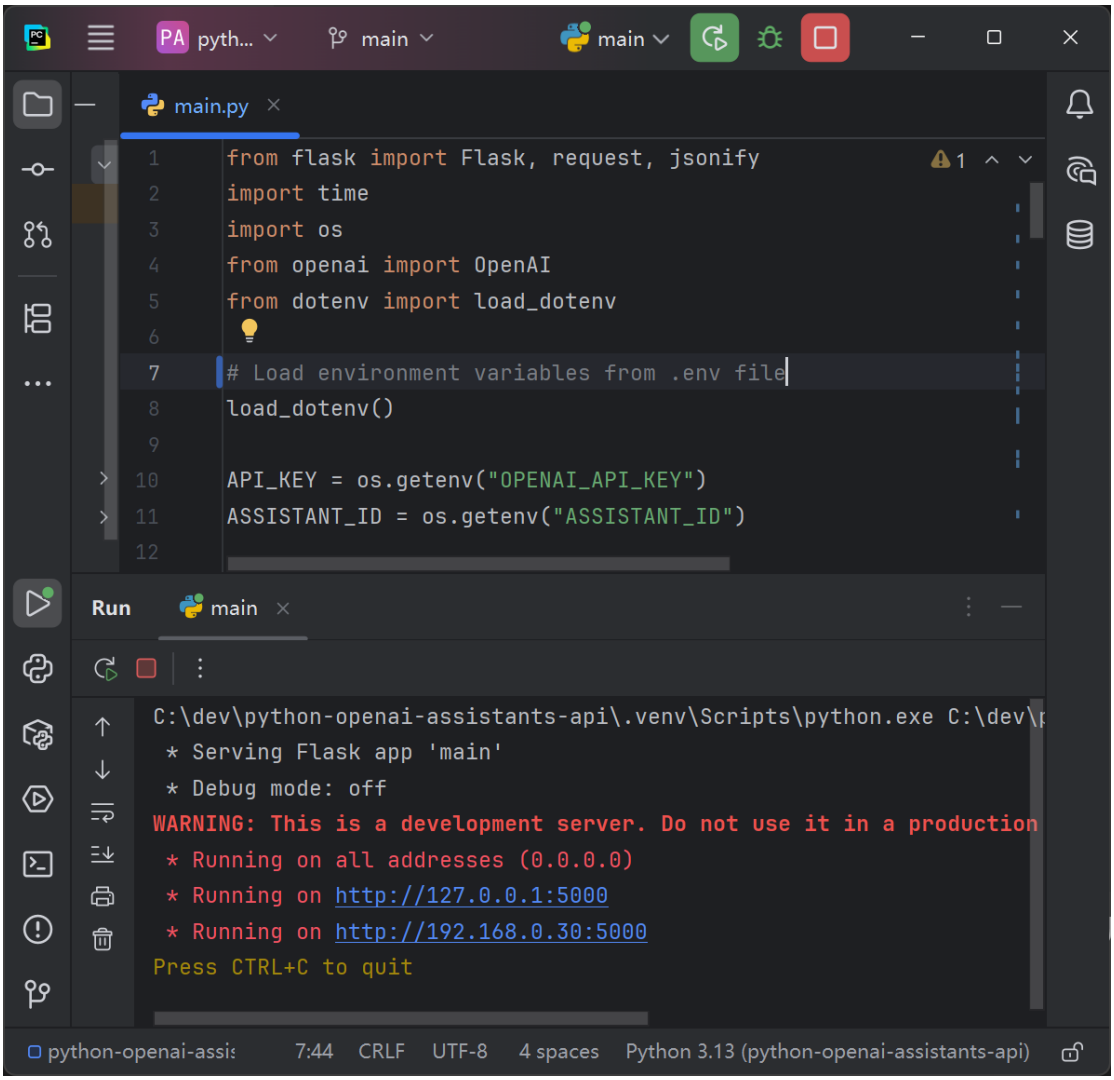


Figure 3.14: Flask-based backend running in PyCharm. The upper panel shows the core assistant logic, including thread management and response polling. The terminal at the bottom confirms the server is running locally and accepting HTTP requests from Unity. This backend must be active before launching the VR system to enable AI response delivery.

model [98]. The transcribed text is sent to the backend and processed by the assistant.

If the embodied avatar is disabled, the AI response is returned as on-screen text only, rendered using Unity's TextMeshPro [99]. When the avatar is active, the same response is passed to Amazon Polly [100], synthesized into audio, and played back in the virtual scene. The lip movement is synchronized using the Oculus LipSync SDK [101], which generates viseme-based facial animations from the audio waveform.

A visual voice activity indicator is shown below the input panel: green when recording is active, red when idle. This indicator helps users confirm recording status at a glance and avoid missed inputs.

To illustrate this interaction mode, Figure 3.15 presents a side-by-side in-game screenshot comparing both AI panel mode (left) and avatar mode (right). In the AI panel mode, the embodied avatar is disabled. The user's transcribed voice input appears in the lower panel, while the assistant's response is displayed above as floating text. A green indicator confirms that voice recording is active, and button hints for sending or recording are clearly visible.

On the right, the avatar mode is activated. Here, the assistant's response is not only shown as text but also delivered through a humanoid virtual teacher, who speaks using synthesized voice generated by Amazon Polly. The avatar provides facial animation and lip-sync, enhancing the sense of human presence during the interaction. Despite the different visual presentations, both modes rely on the same underlying backend and interaction flow.

3.3.4 Embodied Avatar

To simulate a more human-like interaction experience, the system includes an embodied 3D avatar that can be toggled on or off via a "Hide/Active" button. When activated, the avatar is rendered directly in the virtual classroom and delivers spoken responses generated by the AI assistant. This enables a multimodal dialogue interface that complements the text-based AI panel, supporting both auditory and visual engagement.

The avatar was created using Ready Player Me [102], which enables rapid generation of stylized humanoid models that are fully compatible with Unity. Users can customize various features such as facial structure, hairstyle, and clothing. Animations were added using Mixamo [103], including idle postures, expressive gestures, and common response motions, all of which contribute to establishing a more natural and believable visual presence.

In addition to Mixamo, Ready Player Me also provides a limited set of built-in animation presets, such as idle, waving, and blinking [104]. These animations are embedded in the exported GLB files and can be previewed and played directly within the Unity Editor, making them useful for quick prototyping or low-complexity applications.

To support more complex or interactive behaviors, external animations can be imported and applied to the avatar using Unity's built-in Animator system. For humanoid avatars, Unity supports animation retargeting through its Humanoid Rig system, which enables seamless alignment of imported animation clips (such as those from Mixamo) with the avatar's skeletal structure. Detailed implementation guidance can be found in the Unity Animation documentation [105], along with numerous community-authored tutorials.

When the AI assistant returns a text response, the content is simultaneously sent to Amazon Polly [100], which converts it into high-quality speech using Neural Text-

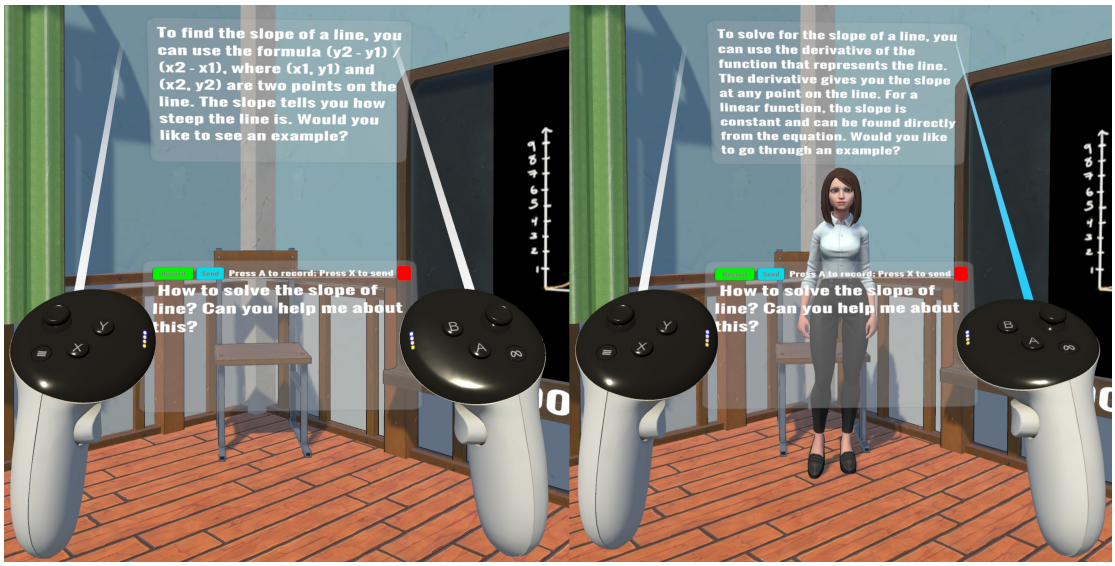


Figure 3.15: Two visual modes of the AI assistant. The left side shows the AI panel mode, in which the embodied avatar is disabled. The assistant's response appears as floating text above the user's transcribed question. The right side shows the avatar mode, where the assistant's answer is delivered via an animated virtual teacher, accompanied by voice synthesis and synchronized lip movements. Both interfaces use the same dialogue backend and support voice-based interaction through the controller buttons. In both modes, users hold Meta Quest controllers to interact with the system. Pressing the A button on the right-hand controller starts voice recording, and pressing it again stops recording. The transcribed query can be sent by pressing the X button on the left-hand controller. As an alternative, users may click the "Record" and "Send" buttons directly on the on-screen panel if using a hand-tracked pointer or prefer not to use physical buttons.

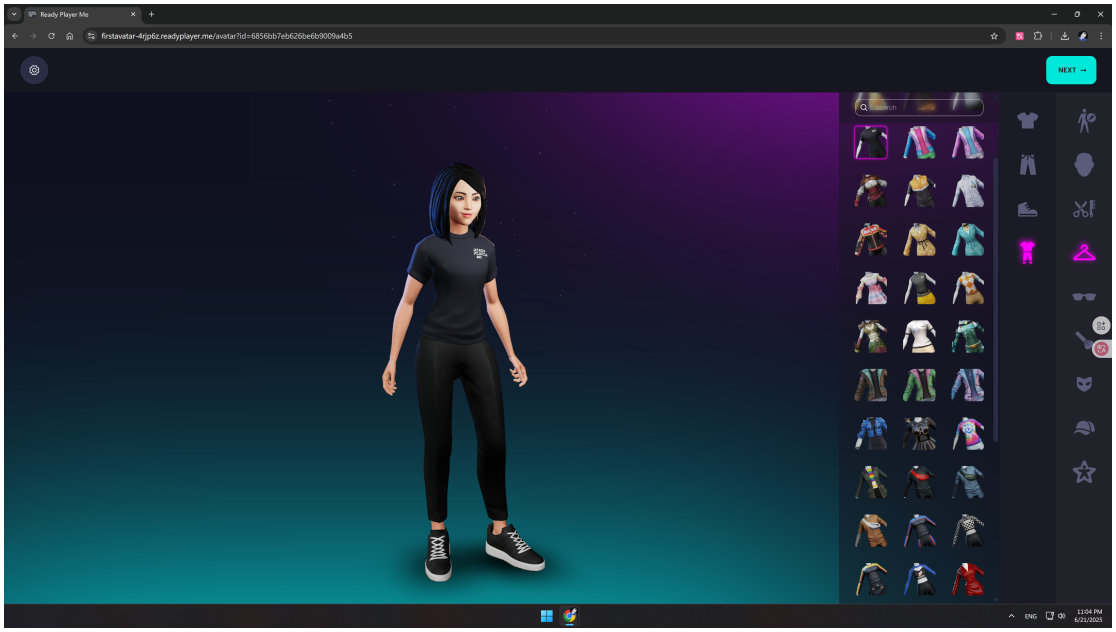


Figure 3.16: Avatar customization interface in Ready Player Me. Users can generate personalized 3D avatars by selecting facial features, hairstyles, outfits, and accessories. The resulting models are stylized, humanoid, and fully compatible with Unity, making them well-suited for integration into immersive applications such as VR classrooms [102]. In this study, a female avatar was created to serve as the embodied AI assistant.

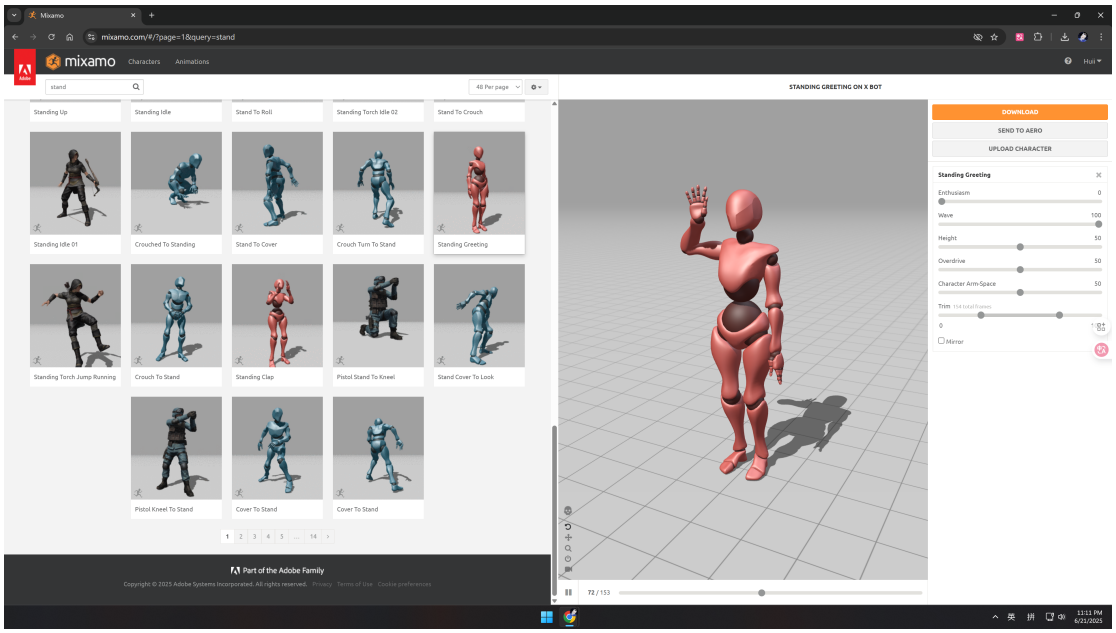


Figure 3.17: The Mixamo animation library interface. Users can search and preview a wide range of pre-made animations, such as gestures, walking cycles, and conversational poses. The selected motion, in this case a standing greeting, is downloadable and can be retargeted onto avatars exported from Ready Player Me. This process enhances the avatar's expressiveness and supports non-verbal communication cues in the VR classroom [103].

to-Speech (NTTS) models. The synthesized audio is then played via Unity's audio playback system and synchronized with the avatar's mouth movement using the Oculus LipSync SDK [101]. This SDK analyzes the audio waveform in real time and applies the corresponding viseme animations to the avatar's mouth mesh, simulating realistic speech articulation.

By combining speech audio, facial motion, and body animation, the avatar functions as a socially present AI tutor within the virtual classroom. This embodiment is designed to enhance user engagement, reduce the perceived distance between the user and the system, and foster a more natural and effective educational interaction.

3.3.5 User Interaction Logging

The system records user interactions locally in two formats: quiz responses and AI conversation logs.

Quiz Answers. At the end of each instructional video, users are prompted to answer a corresponding quiz. Each quiz response is stored in a CSV file containing three columns: `Question`, `Time`, and `Answer`. This structure allows for easy parsing and later statistical analysis. An example entry is shown below:

```
"Quiz 1. What is the average rate of change of g(x) = 2x
over the interval [3, 10]?",
"2025-05-31 15:10:41",
"2"
```

AI Interaction Logs. For voice-based interactions with the AI assistant, the system saves all conversational exchanges in a plaintext log file. Each entry includes the avatar visibility status, the user's input query, the assistant's generated response, and the system's response time. This format allows researchers to trace the interaction context and evaluate AI performance across sessions. An example log is shown below:

```
[2025-06-02 13:31:43] AvatarToggle: Active
[2025-06-02 13:32:32] AvatarToggle: Hide
[2025-06-02 13:49:09] Question: Would you tell me about
this system?
[2025-06-02 13:49:09] Response: Could you please specify
which term or software you're referring to?
[2025-06-02 13:49:09] ResponseTime: 5.24s
```

These logs enable fine-grained analysis of user behavior and system responsiveness. By examining user queries, assistant replies, and timing data, it becomes possible to assess levels of user engagement, conversational coherence, and latency—factors that are critical for evaluating both user experience and system usability in a VR learning context.

Chapter 4

Evaluation

This chapter describes the methods and procedures used to evaluate the proposed VR-based learning system. The evaluation was conducted in two stages: an initial pilot test ($n = 2$) to identify usability and system stability issues, followed by a formal user evaluation study ($n = 18$).

4.1 Pilot Test

Before launching the main user study, we conducted a small-scale pilot test with two male participants to assess the initial usability, technical stability, and perceived user experience of the VR-based system. The participants were selected to represent a range of academic backgrounds and varying familiarity with calculus. One was a first-year undergraduate student majoring in computer science, who had limited exposure to advanced mathematics. The other was a graduate student from a computer science program, with a solid foundation in calculus and prior experience using immersive technologies.

Procedure. Each participant completed a full learning session using the prototype system, which included watching one instructional video (approximately five minutes in length at the time), interacting with the AI assistant through voice input, in the virtual classroom using the Meta Quest 3 headset. We observed their interactions and conducted a short semi-structured interview afterwards to collect open-ended feedback.

Findings. The pilot revealed several usability and interaction issues:

- **Attention and Fatigue:** Both participants reported difficulty maintaining focus during extended video playback in VR. The undergraduate student described feeling "mentally tired" after around three minutes, while the graduate student noted a "drop in concentration" when the content remained passive for too long. These experiences suggest a high cognitive load resulting from prolonged passive viewing in an immersive environment.

- **Navigation Limitations:** Participants expressed frustration at the absence of fast-forward and rewind functions, which made it difficult to revisit unclear segments or skip over known material. This limitation increased perceived cognitive effort and decreased learning efficiency.

- **AI Response Delays:** Although the AI assistant generally provided accurate and relevant responses, both participants noted a delay of approximately 3 to 5 seconds between issuing a query and receiving a spoken reply. This disrupted conversational flow and reduced the sense of real-time interaction.
- **Technical Observations:** No major technical issues, such as crashes or rendering errors, were observed. However, participants mentioned that the voice recording function lacked clear feedback and sometimes captured ambient noise, leading to uncertainty about when the system was actively listening.

Modifications. In response to this feedback, we introduced three primary adjustments before launching the full-scale user study:

1. **Video duration** was shortened to approximately three minutes per segment, reducing fatigue and improving learners' ability to focus on core concepts.
2. **Playback control buttons** (+5 seconds and -5 seconds) were added to the video player, allowing users to easily navigate and rewatch specific segments at their own pace.
3. **Response timing tracking** was implemented to monitor delays in AI replies and support future improvements. Additionally, slight adjustments were made to the assistant's training in the OpenAI Assistants Playground to improve response consistency. However, response latency remains a limitation due to constraints in API processing time, which is currently beyond our direct control.

Conclusion. While the small sample size limits generalizability, the pilot test offered valuable insights that guided meaningful design refinements. It confirmed that the system was technically stable, but highlighted the importance of pacing, feedback clarity, and responsive interactivity to ensure a comfortable and effective learning experience in VR.

4.2 User Study

Building on the insights from the pilot phase, we conducted a full-scale user study to assess the system's effectiveness under real learning scenarios. We adopted a within-subject design in which each participant experienced all three experimental conditions (C1, C2, C3). To counterbalance potential ordering effects, we employed a Latin square arrangement, as shown in Table 4.1. This ensured that each condition appeared equally often in each position and was preceded and followed by every other condition the same number of times.

Although the order of experimental conditions (C1, C2, C3) was counterbalanced using a Latin square, all participants experienced the same fixed sequence of instructional videos (Video 1, Video 2, Video 3) and corresponding quizzes (Quiz 1, 2, 3). That is, the learning content was kept constant across participants, and only the system condition varied per step. As illustrated in Figure 4.1, each participant completed three full sessions, one per condition, before submitting a final condition ranking.

Figure 4.2 shows a real participant engaged in the study while wearing the Meta Quest 3 headset and interacting with the VR classroom environment.

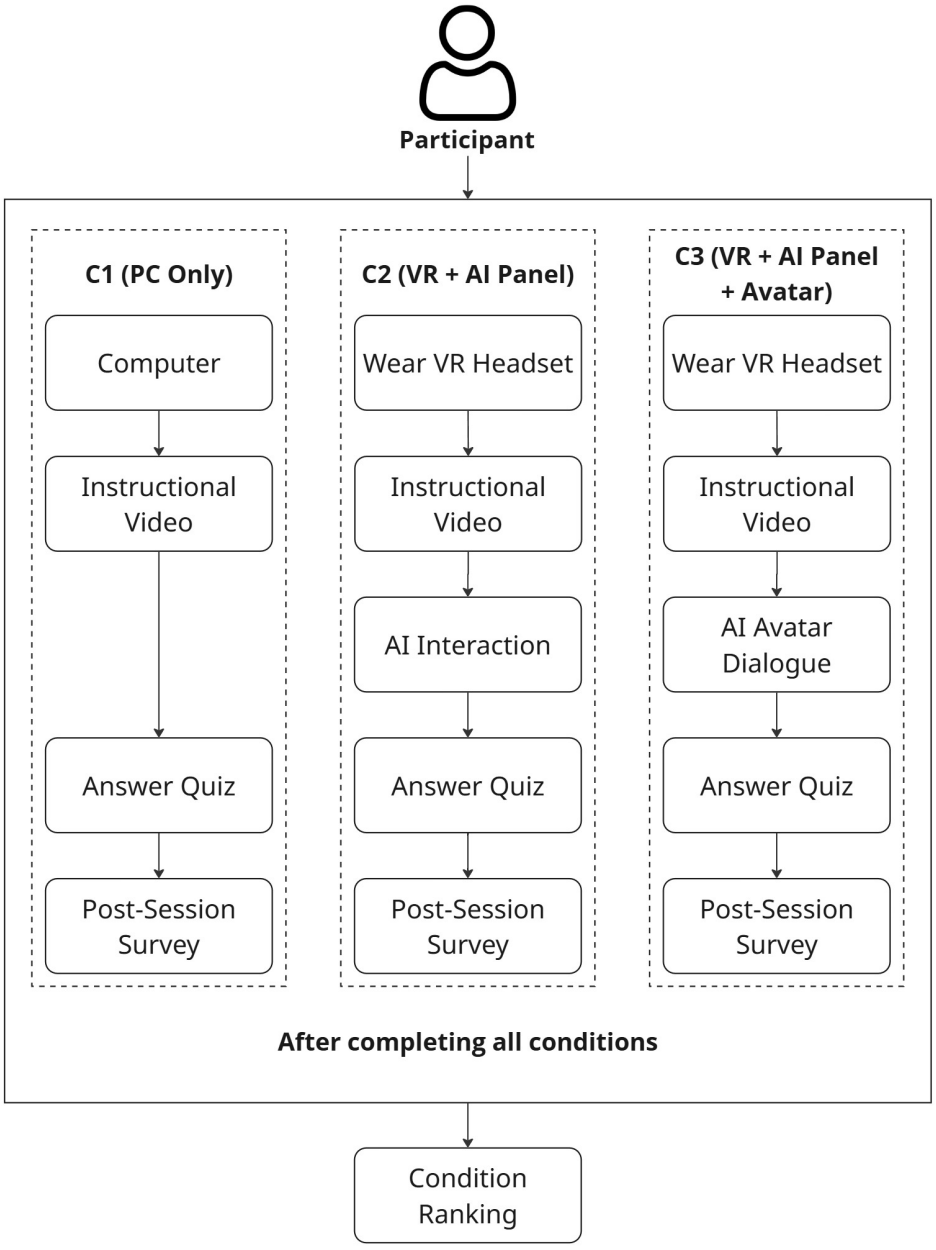


Figure 4.1: Workflow diagram showing participant progression through all three conditions. Each participant experiences C1 (PC only), C2 (VR with AI panel), and C3 (VR with AI avatar), followed by a final ranking task.

Table 4.1: Balanced Latin Square used for participant condition order

Participant ID	Session 1	Session 2	Session 3
P1	C1	C2	C3
P2	C1	C3	C2
P3	C3	C1	C2
P4	C3	C2	C1
P5	C2	C3	C1
P6	C2	C1	C3
P7	C1	C2	C3
P8	C1	C3	C2
P9	C3	C1	C2
P10	C3	C2	C1
P11	C2	C3	C1
P12	C2	C1	C3
P13	C1	C2	C3
P14	C1	C3	C2
P15	C3	C1	C2
P16	C3	C2	C1
P17	C2	C3	C1
P18	C2	C1	C3



Figure 4.2: Participant during the user study. The user is wearing a Meta Quest 3 headset and holding the controllers while watching the instructional video and interacting with the virtual AI assistant. The experiment was conducted in a controlled lab environment, as shown on the connected laptop.

4.2.1 Task

Each participant completed three learning sessions, each corresponding to one of the experimental conditions. In each session, participants watched a 3-minute instructional video on differential calculus and subsequently completed a quiz. In conditions involving AI interaction (C2 and C3), participants were instructed to ask at least one question related to the video content using the speech interface. After each condition, they completed seven subjective ratings related to their experience.

To ensure consistent content across conditions, all participants viewed the same videos in the same order: Video 1 in the first session, Video 2 in the second, and Video 3 in the third.

Before beginning the experimental sessions, participants were asked to complete a short demographic questionnaire via Google Forms. The questionnaire gathered background information including age, gender, prior experience with virtual reality (VR), self-assessed mathematics level and confidence, and English language proficiency. These data were used to describe the participant pool and to explore potential correlations between individual traits and subjective outcomes in the study.

4.2.2 Conditions

The study included the following three conditions:

- **C1:** Baseline condition. Participants watched the video and completed the quiz on a desktop PC without VR or AI assistance. A piece of scratch paper was provided for note-taking and calculation, shown as Fig.4.3.
- **C2:** VR-based video learning with an AI assistant shown via a text-only panel, shown as Fig.4.4.
- **C3:** VR-based video learning with an embodied AI avatar providing spoken responses and lip-synced animation, shown as Fig.4.5.

4.2.3 Measures Overview

To evaluate each condition, we used both standardized and study-specific questionnaire items. Participants completed a custom-designed questionnaire for each learning condition, with all items rated on a 7-point Likert scale ranging from 1 (Strongly Disagree) to 7 (Strongly Agree). To support inclusivity, all items were presented in English, Japanese, and Chinese.

The questionnaire consisted of seven core items administered across all conditions (C1, C2, and C3), along with four additional items presented only in Condition 3, where the AI avatar was available. A summary of all items is provided in Table 4.2.

4.2.4 Adaptation from IPQ and NASA-TLX

This study employed a subset of items adapted from two well-established instruments to measure user experience and cognitive load in the context of self-directed learning using immersive and non-immersive environments. The selected items focused on key constructs most relevant to short educational sessions involving video-based instruction and conversational AI.

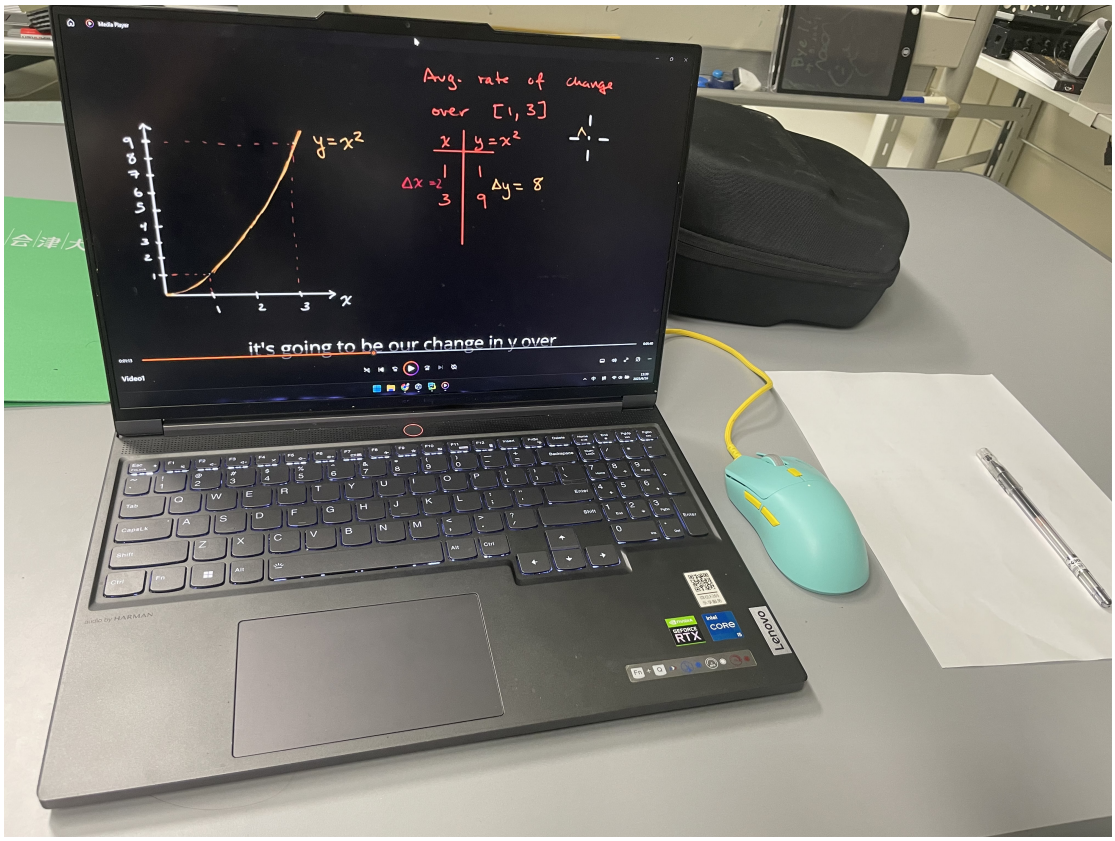


Figure 4.3: Condition 1: PC-based learning setup. The participant watches the video on a standard laptop screen, with a scratch paper and pen for notes or calculation. No VR or AI interaction is used in this condition.



Figure 4.4: Condition 2: VR-based learning with a text-only AI assistant. The participant interacts with the assistant via voice input using the Meta Quest controllers. The AI response appears on a floating panel without any visual avatar. The red ”Hide” toggle button at the top right indicates that the embodied avatar is currently turned off.



Figure 4.5: Condition 3: VR-based learning with a multimodal AI assistant. The participant communicates with a virtual 3D avatar that responds using synthesized speech and lip-synced animation. The avatar stands beside the screen while the user watches and interacts with the video content. The green "Active" toggle button at the top right confirms that the embodied avatar is currently enabled and visible.

IPQ Adaptation. Three core items, namely **Immersion**, **Enjoyment**, and **Comprehension**, were adapted from conceptual dimensions of the Igroup Presence Questionnaire (IPQ) [106]. While the full IPQ includes subscales such as spatial presence and realism, these were omitted due to their limited applicability in brief, structured learning sessions. The selected items were reworded to match the specific educational context and compare both VR and non-VR conditions.

NASA-TLX Adaptation. Four additional items, which include **Mental Demand**, **Physical Demand**, **Effort**, and **Frustration**, were derived from the NASA Task Load Index (NASA-TLX) [107], which is widely used to assess cognitive and emotional workload. The temporal demand and performance dimensions were excluded, as the sessions were self-paced and performance was separately measured through quiz scores. The selected items were designed to capture perceived workload and user stress during the learning tasks across different interface conditions.

4.2.5 Avatar Perception Ratings

In Condition 3, four additional items were introduced to evaluate participants' subjective impressions of the embodied AI avatar. These items formed the Avatar Perception Ratings, a construct designed specifically for this study to assess the avatar's social and aesthetic characteristics. While inspired by dimensions in the AttrakDiff instrument [108], including pragmatic and hedonic quality, the items were newly formulated to better match the avatar's educational role and appearance.

All questionnaire items were written in the first person (e.g., "I felt immersed in the experience") to encourage personal reflection. Responses were collected anonymously using Google Forms immediately after each session, with each form requiring approximately two to three minutes to complete.

Table 4.2: Questionnaire Items Used in Each Condition

Category	**Item**	**Source**	**Condition**
Immersion	I felt immersed in the experience.	IPQ	All
Enjoyment	I enjoyed the experience.	IPQ	All
Comprehension	I understood the content of the experience well.	IPQ	All
Mental Demand	The experience was mentally demanding.	NASA-TLX	All
Physical Demand	The experience was physically demanding.	NASA-TLX	All
Effort	I had to put a lot of effort into this experience.	NASA-TLX	All
Frustration	I felt frustrated during the experience.	NASA-TLX	All
Human-likeness	The AI avatar was human-like.	AttrakDiff-inspired	C3 only
Usefulness	The AI avatar was useful in the experience.	AttrakDiff-inspired	C3 only
Pleasantness	The AI avatar was pleasant.	AttrakDiff-inspired	C3 only
Appearance	I liked the AI avatar's appearance.	AttrakDiff-inspired	C3 only

At the end of the study, participants were also asked to rank the three learning conditions (C1, C2, and C3) in order of personal preference. Rankings were submitted through a dedicated section in the final form and served as an additional indicator of subjective satisfaction. To explore participant rationale, two open-ended questions were included: "Why did you choose to rank this condition as A (most preferred)?" and "Why did you choose to rank this condition as C (least preferred)?"

Each study session included video viewing, optional AI interaction, quiz answering, and questionnaire completion. The entire procedure took between 25 and 45 minutes to complete. Most participants finished within 35 minutes, with only two requiring more time due to extended interaction with the AI assistant.

4.3 User Study Results

4.3.1 Participant Demographics (Pre-questions)

The gender composition (Fig.4.6) of the 18 study participants is presented in the pie chart. The data indicates a non-uniform distribution between the two reported gender categories. The majority group comprised 61.1% of the total sample, while the minority group accounted for the remaining 38.9%.

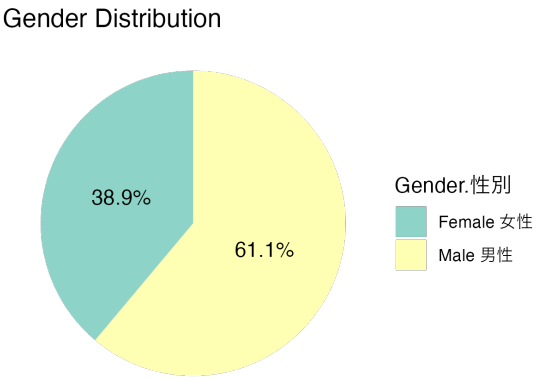


Figure 4.6: Gender distribution of participants showing the proportion of male and female participants in the study.

The age distribution (Fig. 4.7) of the participant cohort is illustrated in the histogram. The sample consisted of 18 individuals with ages ranging from approximately 19 to 33 years. The distribution is unimodal, with the highest frequency of participants (n=6) falling within the 23-to-25-year-old age bracket. The frequencies tail off in both directions from this central peak, suggesting a concentration of participants in their early- to mid-twenties.

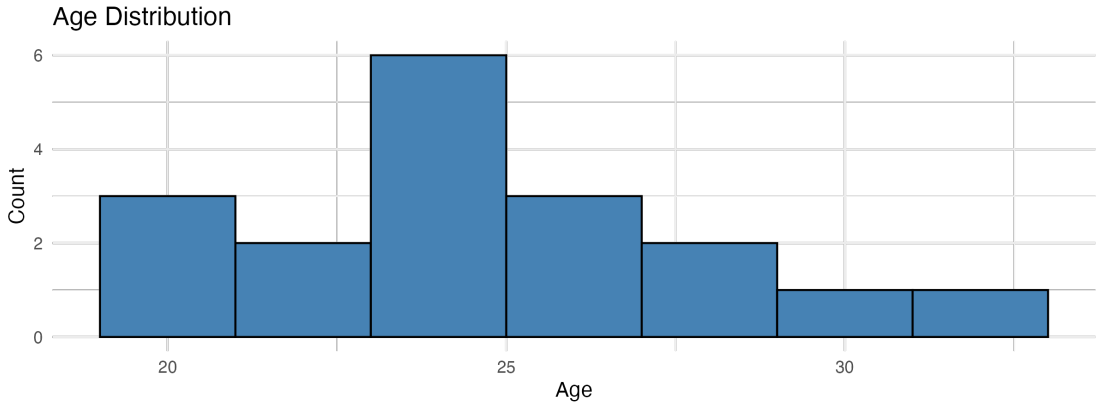


Figure 4.7: Distribution of participant ages showing the frequency of different age groups in the study.

The self-reported mathematics proficiency levels (Fig. 4.8) of the participants shows they were distributed across three distinct levels of mathematical skill. A plurality of participants (44.4%) identified with the most common proficiency level (represented in blue), followed by 33.3% in the second category (green) and 22.2% in the third (red).

This distribution highlights a varied mathematical background among the participants, with no single group forming a majority.

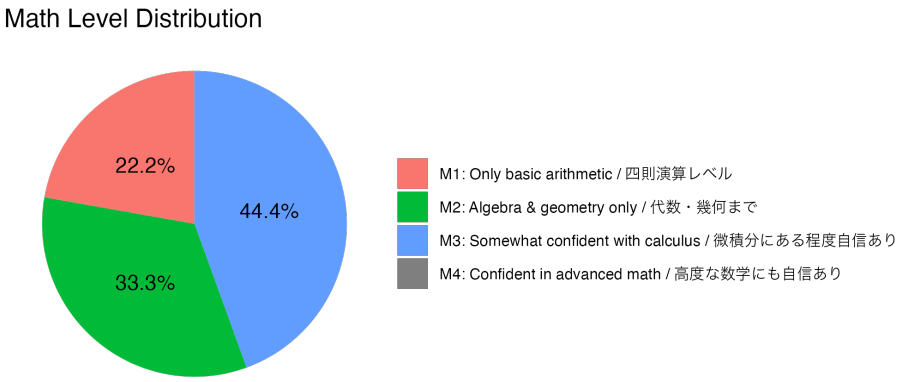


Figure 4.8: Distribution of mathematical proficiency levels among participants, showing the proportion of different math skill levels.

The distribution of self-assessed English language proficiency (Fig. 4.9) among the participants is displayed in the pie chart. The responses were spread across four categories. The largest group, accounting for 44.4% of participants, reported the proficiency level indicated in yellow. The next most common level was reported by 22.2% of the sample (teal), while the remaining two levels (purple and red) were each reported by 16.7% of participants.

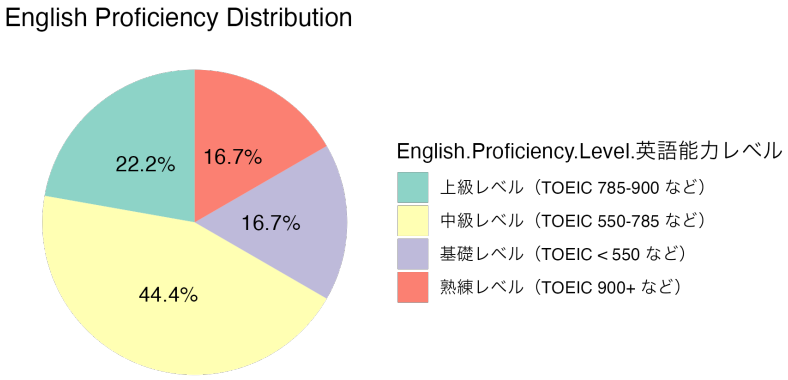


Figure 4.9: Distribution of English proficiency levels among participants, showing the proportion of different language skill levels.

The participants' prior experience with virtual reality (VR) (Fig. 4.10 shows significant majority of participants (55.6%) reported having the level of VR experience denoted by the yellow slice. The remaining participants were distributed among four other categories, with two groups each constituting 16.7% of the sample (pink and teal) and two smaller groups each representing 5.6% (blue and red). This indicates that over half the cohort shared a common background in VR usage, while the rest possessed a range of different experience levels.

The distribution of virtual reality experience levels for distinct subgroups (Fig. 4.11) based on their mathematical proficiency (M1, M2, M3) reveals that participants in the M1 proficiency group reported experience primarily in two VR categories (75.0% and

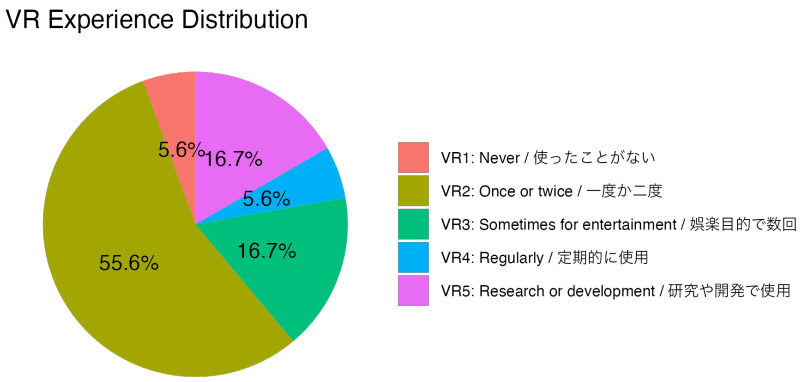


Figure 4.10: Distribution of VR experience levels among participants, showing the proportion of different VR usage patterns.

25.0%). The M2 group showed slightly more diversity in VR experience. Notably, the M3 group exhibited the greatest heterogeneity, with their prior VR experience spread across four different categories. No data were available for proficiency levels M4 and M5.

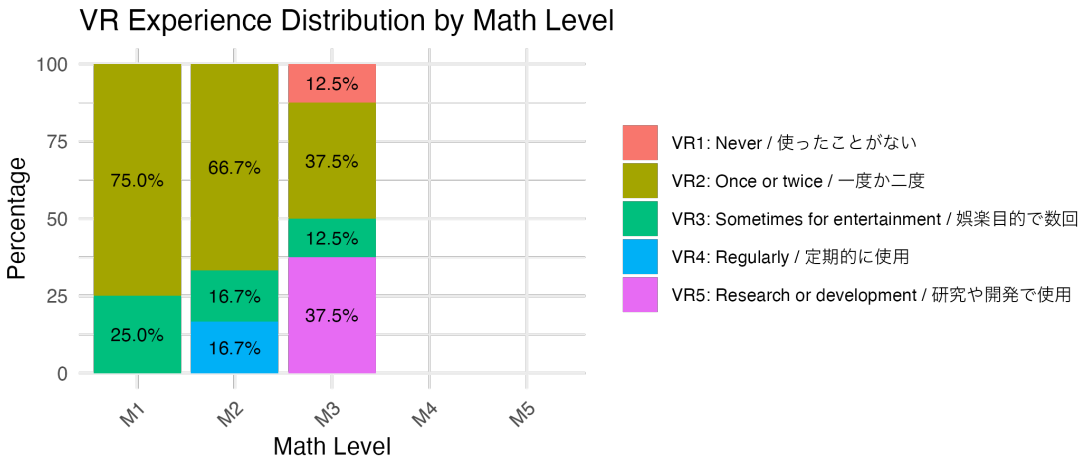


Figure 4.11: Distribution of VR experience levels across different mathematical proficiency categories, showing the relationship between math skills and VR experience.

The composition of mathematics proficiency for each of the five VR experience groups (VR1–VR5) (Fig. 4.12) indicates that certain VR experience groups were homogenous in terms of mathematical background; for instance, the VR1 and VR5 groups were composed entirely of participants from a single proficiency level (represented by the dark green category), and the VR4 group was composed entirely of participants from another level (dark yellow). In contrast, the VR2 and VR3 experience groups were heterogeneous, comprising participants from three distinct math proficiency levels in varying proportions.

4.3.2 IPQ and NASA-TLX Results

This study employed a statistical approach to analyze questionnaire responses across three experimental conditions. The analysis utilized cumulative link mixed models

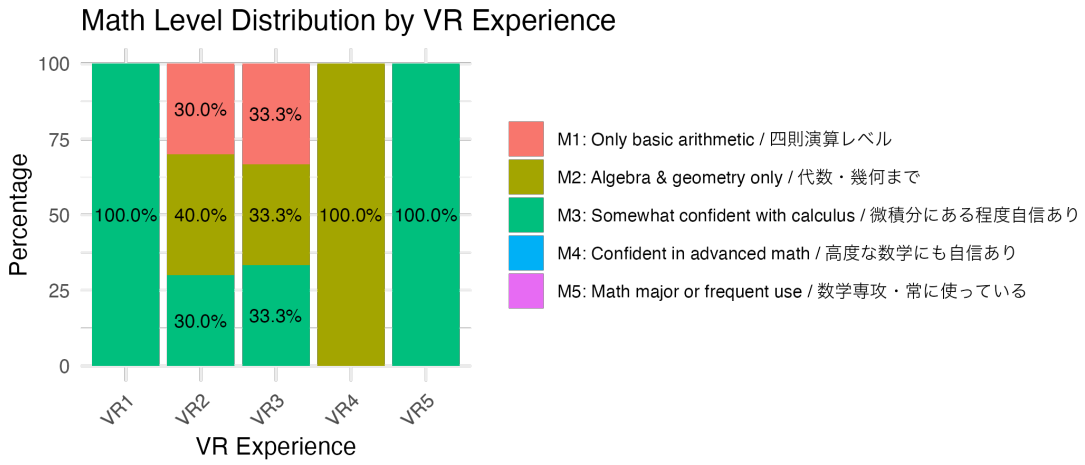


Figure 4.12: Distribution of mathematical proficiency levels across different VR experience categories, showing the relationship between VR experience and math skills.

(CLMM) to account for the ordinal nature of the response data while incorporating random effects for individual participants. Seven key measures were examined: Immersion, Enjoyment, Comprehension, Mental Demand, Physical Demand, Effort, and Frustration. For each measure, the model included condition as a fixed effect and participant as a random effect, allowing for the examination of condition effects while accounting for individual differences. The statistical analysis was conducted using the ordinal package in R. Post-hoc analyses were performed using estimated marginal means (EMM) with pairwise comparisons to identify specific differences between conditions. The results were visualized using ggplot2, presenting estimated marginal means with 95% ($p < 0.05$) confidence intervals for each condition.

In the context of CLMM analysis, the *estimate* is the coefficient difference between two conditions in the logistic regression framework. Since this data using ordinal regression with a logit link function, the estimate represents how much the log-odds of being in a higher category changes when moving from one condition to another. *SE* stands for Standard Error, which is a measure of the precision or uncertainty of the estimate. The SE tells how much the estimate might vary if the same experiment is repeated multiple times with different samples from the same population.

The analysis of **immersion** (Fig. 4.13) ratings revealed a highly significant main effect of condition ($\chi^2(2) = 18.25, p < 0.001$). Post-hoc pairwise comparisons demonstrated that participants reported significantly lower immersion in condition C1 compared to both C2 (estimate $= -2.64$, SE $= 0.78, p = 0.002$) and C3 (estimate $= -3.79$, SE $= 0.90, p < 0.001$). The comparison between C2 and C3 was not statistically significant (estimate $= -1.15$, SE $= 0.66, p = 0.188$), indicating that both conditions C2 and C3 provided similarly enhanced immersive experiences relative to the baseline condition C1.

Enjoyment (Fig. 4.14) ratings showed a significant condition effect ($\chi^2(2) = 14.57, p < 0.001$), following a similar pattern to immersion. Pairwise comparisons revealed that condition C1 yielded significantly lower enjoyment ratings compared to both C2 (estimate $= -2.35$, SE $= 0.77, p = 0.006$) and C3 (estimate $= -3.19$, SE $= 0.85, p < 0.001$). No significant difference was observed between conditions C2 and C3 (estimate $= -0.84$, SE $= 0.68, p = 0.431$), suggesting that both experimental conditions provided comparable levels of user enjoyment that were superior to the baseline

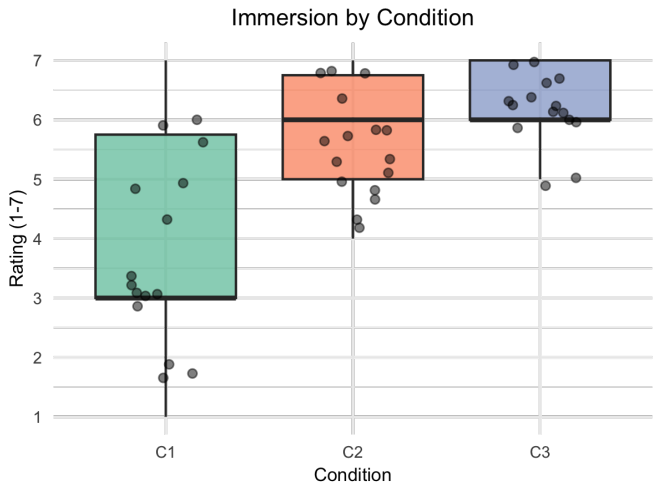


Figure 4.13: Immersion

condition.

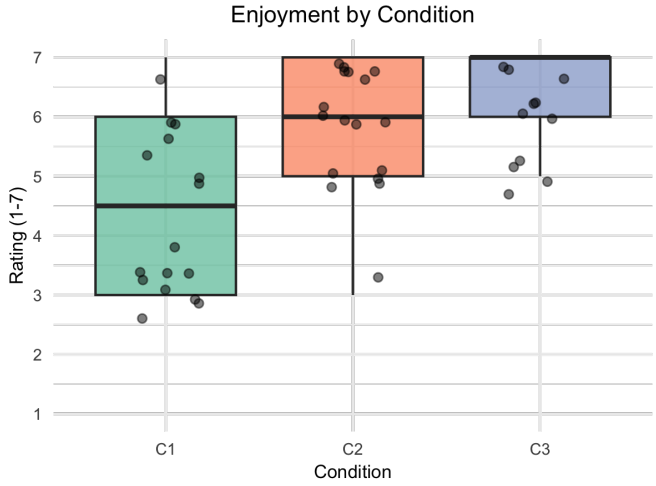


Figure 4.14: Enjoyment

The analysis of **comprehension** (Fig. 4.15) ratings did not reveal a significant main effect of condition ($\chi^2(2) = 2.50, p = 0.287$). Pairwise comparisons confirmed the absence of significant differences between any condition pairs: C1 vs C2 (estimate $= -0.79$, SE $= 0.66, p = 0.456$), C1 vs C3 (estimate $= -1.04$, SE $= 0.69, p = 0.286$), and C2 vs C3 (estimate $= -0.25$, SE $= 0.66, p = 0.925$). These results indicate that participants' perceived comprehension remained consistent across all experimental conditions, suggesting that the different conditions did not differentially impact cognitive understanding or clarity.

Mental demand (Fig. 4.16) ratings showed no significant condition effect ($\chi^2(2) = 1.77, p = 0.413$). All pairwise comparisons were non-significant: C1 vs C2 (estimate $= 0.72$, SE $= 0.68, p = 0.535$), C1 vs C3 (estimate $= 0.86$, SE $= 0.69, p = 0.428$), and C2 vs C3 (estimate $= 0.14$, SE $= 0.66, p = 0.977$). The results suggest that the cognitive workload required by participants remained equivalent across all conditions, indicating that the experimental manipulations did not impose differential mental processing demands.

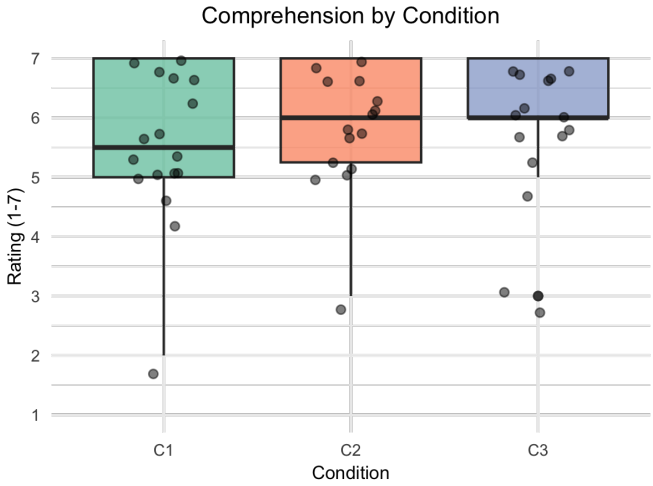


Figure 4.15: Comprehension

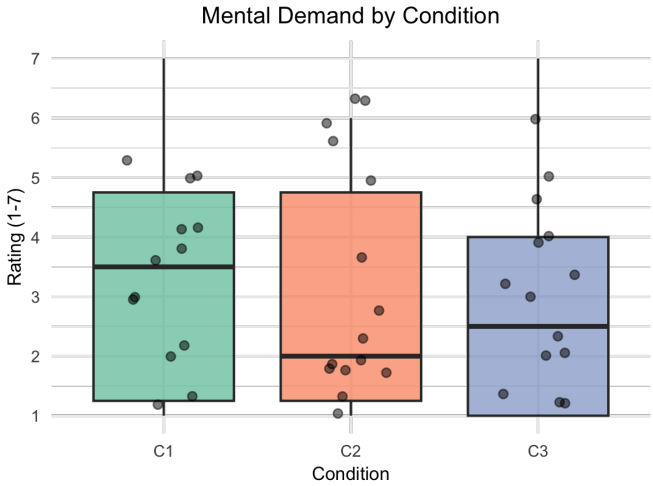


Figure 4.16: Mental Demand

Physical demand (Fig. 4.17) ratings demonstrated no significant main effect of condition ($\chi^2(2) = 0.66, p = 0.717$). Pairwise comparisons were all non-significant: C1 vs C2 (estimate $= 0.40$, SE $= 0.70, p = 0.834$), C1 vs C3 (estimate $= -0.15$, SE $= 0.70, p = 0.975$), and C2 vs C3 (estimate $= -0.55$, SE $= 0.70, p = 0.710$). These findings indicate that participants perceived similar levels of physical effort and bodily demands across all experimental conditions, suggesting that the conditions were equivalent in terms of physical workload requirements.

The analysis of **Effort** (Fig. 4.18) ratings revealed no significant condition effect ($\chi^2(2) = 2.19, p = 0.334$). None of the pairwise comparisons reached statistical significance: C1 vs C2 (estimate $= 0.74$, SE $= 0.65, p = 0.490$), C1 vs C3 (estimate $= 0.92$, SE $= 0.65, p = 0.330$), and C2 vs C3 (estimate $= 0.18$, SE $= 0.61, p = 0.953$). The results suggest that participants invested comparable levels of overall effort across all conditions, indicating that the experimental manipulations did not differentially influence the perceived exertion required to complete the tasks.

Frustration (Fig. 4.19) ratings showed no significant main effect of condition ($\chi^2(2) = 0.11, p = 0.948$), representing the weakest condition effect among all measured variables. All pairwise comparisons were non-significant: C1 vs C2 (estimate $=$

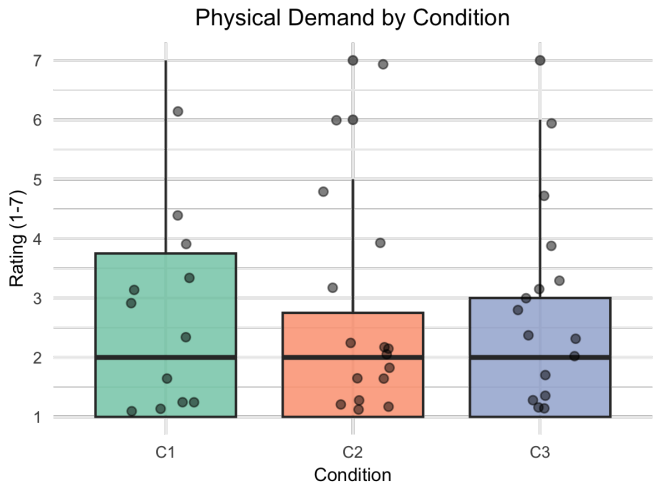


Figure 4.17: Physical Demand

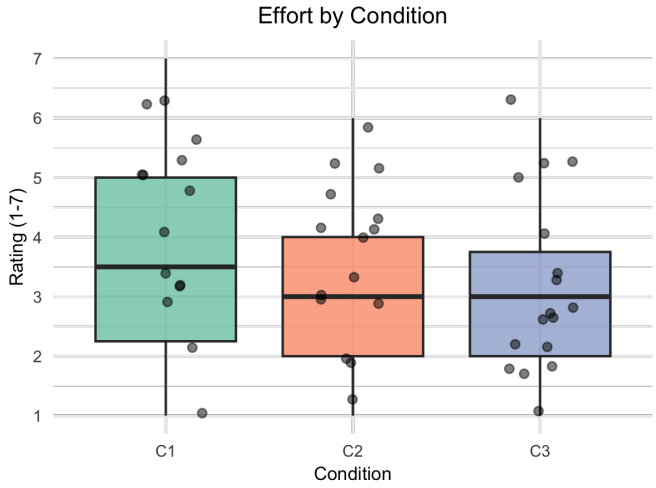


Figure 4.18: Effort

$0.08, \text{SE} = 0.70, p = 0.993$), C1 vs C3 (estimate $= -0.14, \text{SE} = 0.70, p = 0.979$), and C2 vs C3 (estimate $= -0.22, \text{SE} = 0.67, p = 0.944$). These results indicate that participants experienced equivalent levels of frustration across all experimental conditions, suggesting that none of the conditions induced differential negative emotional responses or task-related irritation.

Overall (Fig. 4.20), the analysis revealed significant effects of condition on *Immersion* and *Enjoyment*, with C3 (embodied avatar) outperforming both C1 and C2. No significant differences were observed in *Comprehension* or cognitive load measures (*Mental Demand*, *Physical Demand*). While the differences in *Effort* and *Frustration* were not statistically significant, participants reported lower scores for these items in AI-assisted conditions (C2 and C3), indicating that the presence of an AI assistant may reduce perceived difficulty and enhance user experience.

4.3.3 Quiz Answers

Fig. 4.21 illustrates the mean quiz scores achieved by participants across the three experimental conditions, labeled C1, C2, and C3. The y-axis represents the average

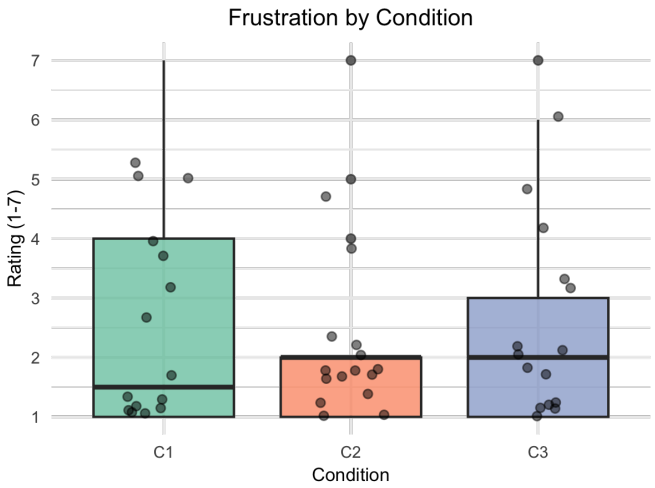


Figure 4.19: Frustration

Question	χ^2	df	p-value
Immersion	18.25	2	0.000109
Enjoyment	14.57	2	0.000686
Comprehension	2.5	2	0.287
Mental Demand	1.77	2	0.413
Physical Demand	0.66	2	0.717
Effort	2.19	2	0.334
Frustration	0.11	2	0.948

Table 4.3: Main effects of condition on each question

score, presumably on a scale of 0 to 100. Condition C1 yielded a mean score of approximately 78. Conditions C2 and C3 resulted in higher and nearly identical mean scores, both registering at approximately 83. These results suggest that participants performed better on the quiz when under conditions C2 and C3 compared to condition C1.

Fig. 4.22 presents the individual quiz scores for each of the 18 participants (P1 through P18). The y-axis represents the score achieved, on a 3-point scale. The data reveal significant variation in performance among individuals. A total of nine participants (P1, P5, P10, P11, P12, P13, P14, P15, P16) achieved the maximum possible score of 3. Conversely, three participants (P3, P7, P9) received a score of 1, the lowest score recorded. The remaining six participants scored 2 points. This distribution highlights the diverse range of outcomes at the individual level.

4.3.4 Ranking Results (post-questions)

Figure 12: Ranking Preference Frequencies by Condition This 100% stacked bar chart displays the frequency with which each condition (C1, C2, C3) was ranked by the participants. The legend indicates three ranking positions (e.g., 1st, 2nd, 3rd choice). For condition C1, a majority (approximately 60%) of participants assigned it their lowest rank (blue), with the remainder split between the first and second ranks. Condition C2 was most frequently assigned the middle rank (orange), by roughly 55% of participants. Condition C3 was predominantly selected as the top-ranked preference (teal),

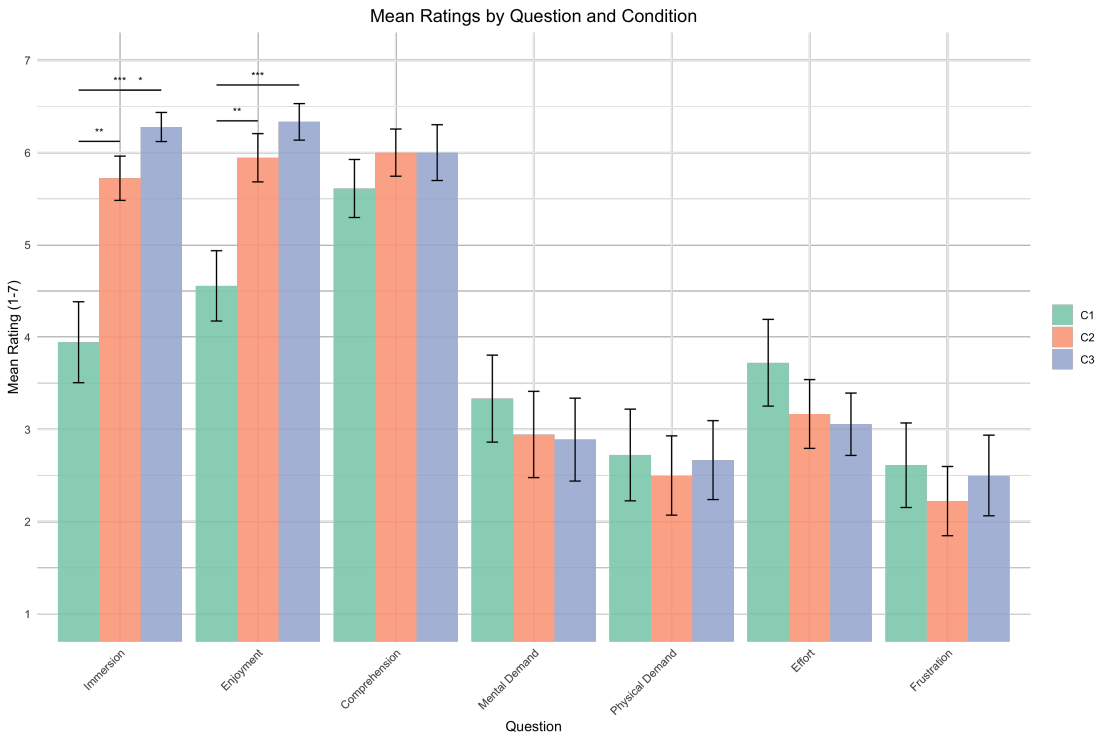


Figure 4.20: Combined bar plot showing mean ratings for all questions across conditions. Significant pairwise comparisons shown in Immersion and Enjoyment questions.

Question	Comparison	Estimate	SE	p-value
Immersion	C1 vs C2	-2.64	0.78	0.00197
Immersion	C1 vs C3	-3.79	0.9	7.06e-05
Enjoyment	C1 vs C2	-2.35	0.77	0.00644
Enjoyment	C1 vs C3	-3.19	0.85	0.000553

Table 4.4: Significant pairwise comparisons between conditions

with approximately 73% of participants assigning it their highest rank. This demonstrates a clear preference order of $C3 > C2 > C1$ among the cohort.

Figure 13: Statistical Summary of Condition Rankings The boxplot provides a statistical summary of the rankings for the three conditions, C1, C2, and C3. The y-axis represents the rank, where a lower value indicates a more preferable rank (e.g., 1 is better than 3). Condition C3 was the most preferred, with a median ranking of 1. Conversely, condition C1 was the least preferred, with a median ranking of 3. Condition C2 occupied the middle position with a median ranking of 2. The clear separation of the boxes, with no overlap in their interquartile ranges, indicates a statistically significant and consistent preference ranking among the participants for C3 as the best, C2 as the intermediate, and C1 as the least preferred condition.

4.3.5 Avatar Perception Ratings

Fig 4.25 shows statistical summary of participant ratings for the virtual avatar across four perceptual dimensions: Appearance, Human-likeness, Pleasantness, and Usefulness, based on a 7-point scale. ”Pleasantness” achieved the highest median rating (ap-

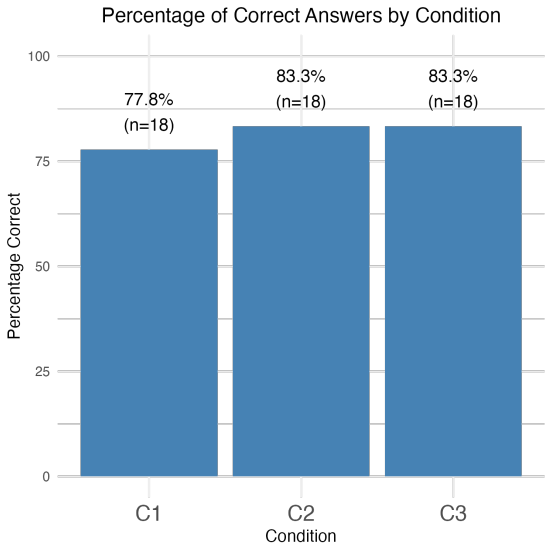


Figure 4.21: Mean quiz scores across three experimental conditions. The bars represent the average score achieved by participants in conditions C1, C2, and C3.

prox. 6.5), followed closely by ”Appearance” and ”Usefulness” (both approx. 6.0). ”Human-likeness” received a notably lower median rating of approximately 5.0 and exhibited the greatest variance in scores, indicating a lack of consensus among participants on this particular attribute. The ratings for ”Pleasantness” were the most consistently high, showing the smallest inter-quartile range.

The heatmap (Fig. 4.26) provides a disaggregated view of the avatar perception ratings for each of the 18 participants (P1–P18) across the four metrics. Lighter shades correspond to higher ratings. The results corroborate the aggregated findings, with ”Pleasantness” and ”Appearance” showing predominantly light coloration, confirming their high scores. Conversely, the ”Human-likeness” column displays significant color variation, visually representing the high inter-participant disagreement previously noted in the boxplot. This visualization also allows for the identification of individual rating patterns, such as certain participants (e.g., P7) who consistently rated the avatar lower across multiple dimensions compared to their peers.

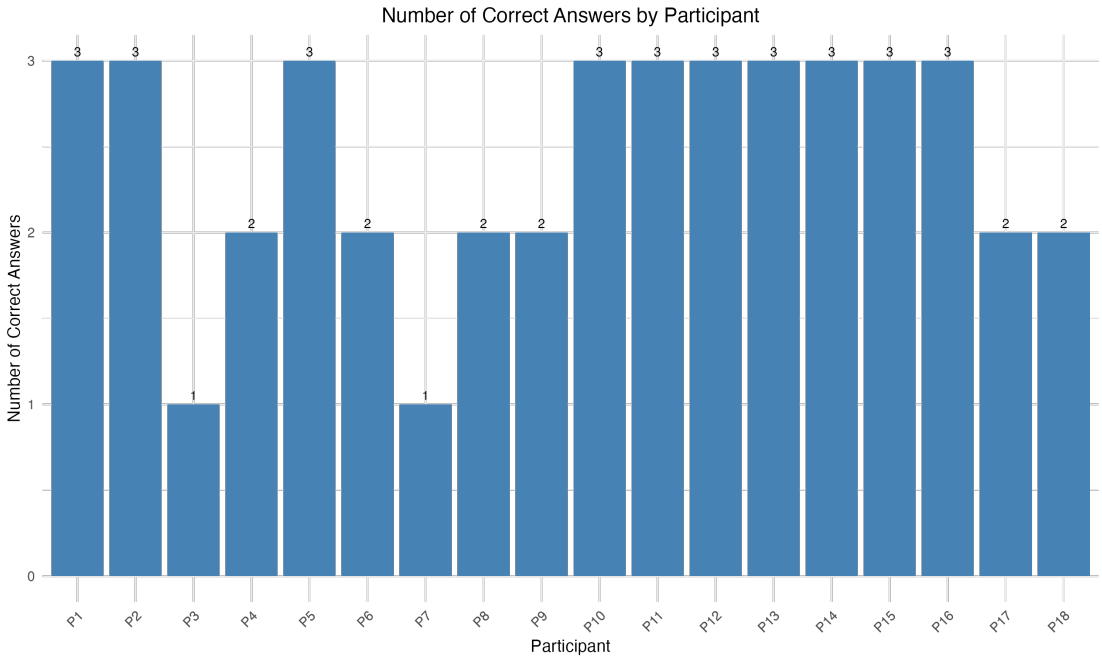


Figure 4.22: Individual quiz scores for each of the 18 participants. The height of the bar corresponds to the score achieved on a 3-point scale.

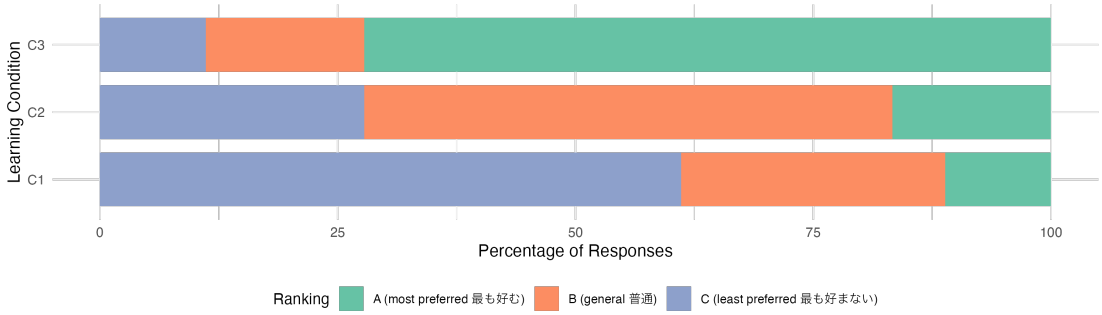


Figure 4.23: Frequency distribution of rankings for the three experimental conditions. The chart illustrates the percentage of participants who ranked each condition (C1, C2, C3) as their first, second, or third preference

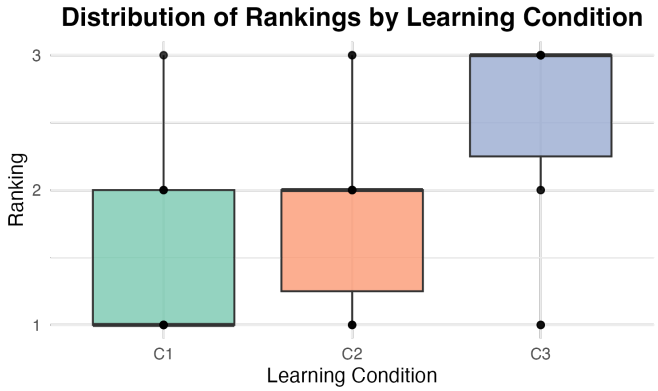


Figure 4.24: Boxplot of participant rankings for the three conditions. The plot displays the median and inter-quartile range of ranks assigned to conditions C1, C2, and C3, where a lower numerical value indicates a lower preference (3=highest, 1=lowest)

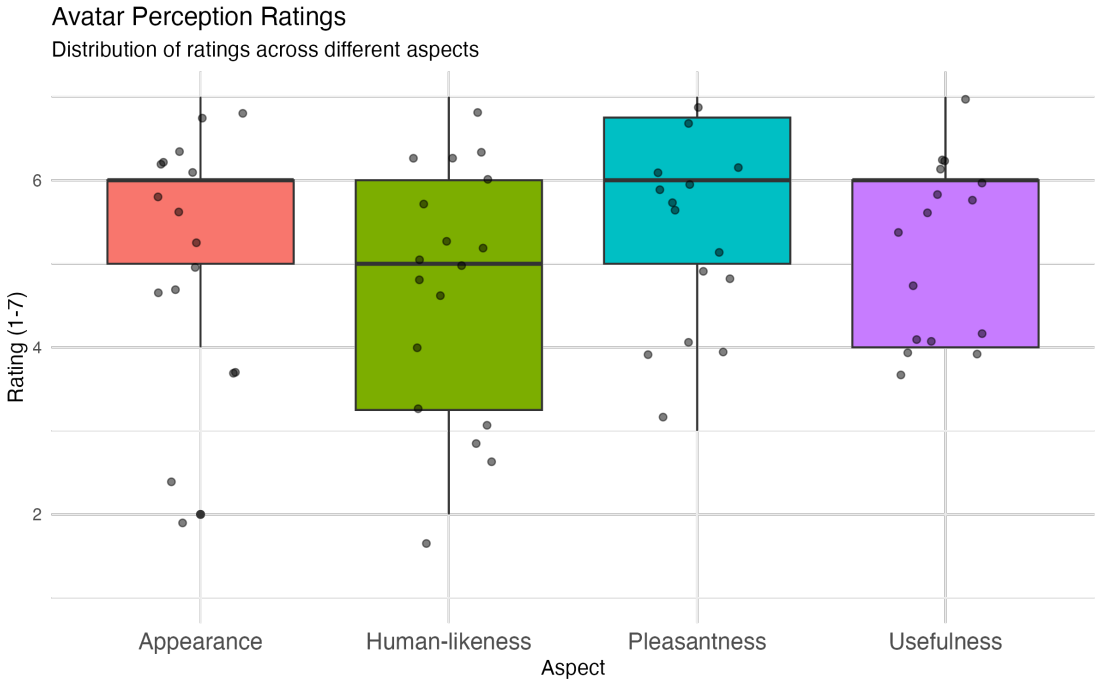


Figure 4.25: Distribution of ratings across different aspects of avatar perception. The box plot shows the median (middle line), quartiles (box boundaries), and individual data points (dots) for each aspect. The aspects evaluated are Human-likeness, Usefulness , Pleasantness, and Appearance. The ratings are on a scale of 1-7, where higher values indicate more positive perceptions.

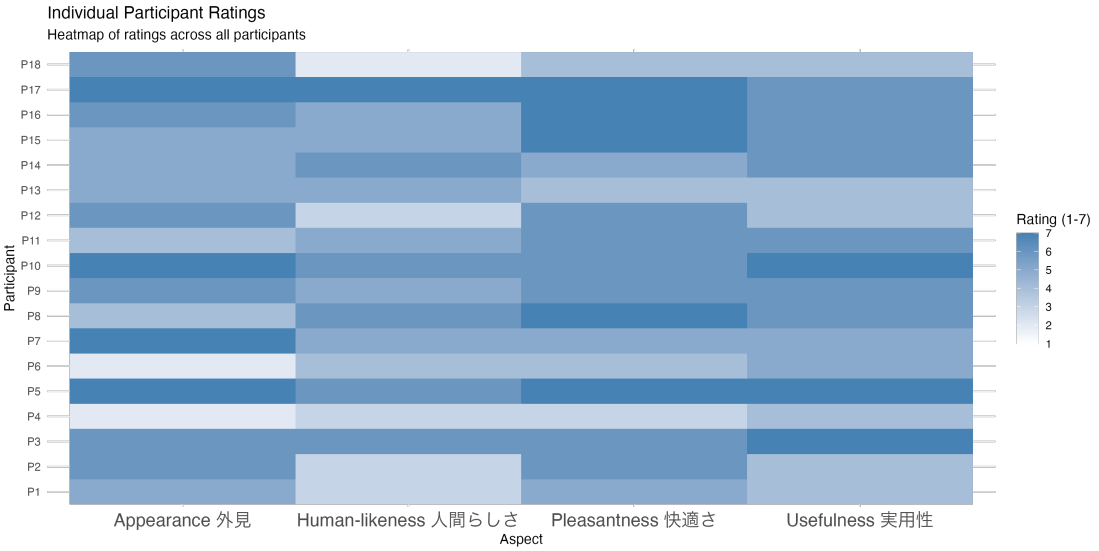


Figure 4.26: Heatmap showing individual participant ratings across all aspects. Each cell represents a single rating from a participant (P1-P18) for a specific aspect. The color intensity indicates the rating value, with darker blue representing higher ratings. This visualization helps identify patterns in individual responses and potential outliers or consistent patterns across participants.

Chapter 5

Discussion

Building upon the empirical findings presented in Chapter 4, this chapter revisits the research questions and hypotheses and interprets the user study results in light of these original goals.

RQ1 asked whether the addition of an embodied avatar to a generative AI assistant would enhance learners' engagement and enjoyment. The results partially support this: while participants ranked the avatar condition (C3) highest in preference and reported slightly higher enjoyment than in the text-only AI condition (C2), the difference was not statistically significant. This suggests that embodied avatars may offer affective advantages, but further refinement is needed to fully realize their potential. Thus, **H1** was not strongly supported, but qualitative feedback suggests promising directions for improvement.

RQ2 investigated how VR-based learning (C2 and C3) compares to traditional PC-based learning (C1) in terms of immersion, enjoyment, and performance. The findings clearly support this hypothesis for subjective measures: both VR conditions yielded significantly higher immersion and enjoyment scores than the desktop baseline, aligning with **H2**. However, quiz performance did not differ significantly between conditions, indicating that while VR enhanced the learning experience emotionally, it did not yet yield measurable gains in short-term comprehension. This may be due to the limited session duration, video-centered instruction, or the simplicity of quiz items.

The following sections analyze these findings more deeply, including system limitations, user preferences, and areas for future development.

5.1 Limitations

Several limitations should be considered when interpreting the findings of this study.

First, the sample size was relatively small (n=18), which may limit the generalizability of results. While within-subject designs help reduce inter-participant variability, a larger sample would enable more robust statistical conclusions and subgroup analyses.

Second, the language used in the system—including videos and AI responses—was English-only. This may have hindered comprehension for some participants, particularly non-native speakers, and potentially influenced their engagement and perceived learning outcomes.

Third, the AI assistant lacked awareness of the video content that had just been played. This disconnect occasionally resulted in off-topic or generic responses, weakening the intended dialogic learning experience.

Fourth, speech recognition was not always reliable. Some participants experienced difficulties having their responses accurately transcribed, which may have disrupted the flow of interaction.

Finally, the avatar design was limited in animation and interactivity. Its static posture and lack of synchronized visual cues were mentioned by participants as reducing the perceived realism and responsiveness of the AI agent.

5.2 Future Work

Based on study results and participant feedback, several directions for future improvement have emerged:

- **Tighter AI-video integration:** Enabling the AI assistant to recognize and reference the instructional video just watched would allow for more coherent, context-aware interactions.
- **Multilingual support:** Incorporating speech input/output in multiple languages and offering subtitles or adjustable speech speed could expand accessibility for diverse learners.
- **Enhanced avatar interactivity:** Improving avatar animation, adding gaze behaviors or natural gestures, and enabling it to initiate conversation could increase perceived presence and engagement.
- **Interactive learning tools:** Introducing in-VR note-taking features, adaptive quizzes, or interactive whiteboards could enhance active learning and support deeper user agency.
- **Scalability across domains:** Extending the system to cover a wider range of subjects, user profiles, and learning tasks would support its broader application as an AI-enhanced education platform.

5.3 User Feedback

5.3.1 Immersion and Classroom Realism

A majority of participants (5 out of 18) described the VR classroom as ”realistic” and ”immersive,” noting that it felt similar to a physical classroom. This sense of presence is consistent with existing literature suggesting that spatial environments can enhance cognitive engagement in self-learning. Additionally, 3 participants noted that the VR environment helped reduce external distractions, especially compared to learning via traditional desktop setups. This suggests VR may be particularly beneficial for participants who struggle with focus in online learning contexts.

5.3.2 Lack of Interaction in the VR Setting

Despite the immersive environment, 5 participants also expressed that the VR classroom felt ”empty” or ”lacked interaction.” While visual and spatial cues were appreciated, the absence of dynamic content or responsive activities limited engagement. This

feedback indicates a need to enrich the VR classroom with more interactive features: such as AI-initiated prompts, ambient responses, or adaptive feedback mechanisms, to simulate a more engaging classroom experience.

5.3.3 Perceived Potential and Limitations of the AI Assistant

Three participants explicitly commented that the AI character seemed promising and ”could be very helpful,” but current limitations undermined its impact. In particular, participants noticed that the AI was not contextually aware of which video had just been watched, reducing the relevance and coherence of its answers. Furthermore, 2 participants noted that voice recognition occasionally failed to register speech input accurately. These issues point to two key areas for improvement: tighter integration between the video module and the AI assistant, and enhanced robustness in speech input handling.

5.3.4 Challenges for Non-native English Speakers

A recurring concern among non-native English participants (3 mentions) was that the video content was delivered too quickly, making comprehension difficult. In addition, the AI assistant supported only English responses, which further limited accessibility. These findings highlight the importance of multi-language support and adjustable video playback speed to accommodate a broader learner demographic.

5.3.5 Avatar Design: Presence and Realism

Participants expressed mixed responses regarding the avatar. While some (3 participants) wished it were “more human-like,” others (2 participants) commented positively on its potential to increase focus and engagement. However, 2 participants also described awkwardness when the avatar stood silently at a distance or remained static throughout the session. Suggestions included enabling the avatar to proactively initiate interaction, synchronize with video content, or visually respond during conversations. These insights point toward the need for improved behavioral realism and contextual responsiveness in embodied agents.

5.3.6 Quiz Simplicity and Task Relevance

Two participants noted that the quizzes were too easy and did not require asking questions to the AI assistant. This observation suggests a mismatch between the difficulty of learning materials and the designed affordances of AI interaction. Future versions could consider adaptive quizzes or open-ended tasks that better leverage the AI assistant’s dialogic capabilities.

5.3.7 Note-taking Support in VR Environments

Three participants suggested adding note-taking functionality to the VR environment, such as a virtual notebook or scratchpad for calculations. In the desktop condition (C1), participants had access to physical paper, which facilitated active engagement. However, this support was absent in VR, leading to a perceived gap in usability and

learning flow. Incorporating digital notetaking tools: such as spatial sticky notes, virtual whiteboards, or handwriting panels, could enhance learning-by-doing and maintain immersion without sacrificing functionality.

5.3.8 Implications for System Scalability

Beyond core feedback, participants offered valuable suggestions for future functionality: support for additional languages, more personalized avatars, and deeper content linkage between instructional materials and AI dialogue. These enhancements align with the goals of creating scalable, inclusive, and engaging learning platforms.

Chapter 6

Conclusion

This thesis explored the integration of embodied avatars and generative AI in virtual reality (VR) to support self-directed learning in advanced mathematics education. The system developed consisted of three core components: (1) short instructional video modules, (2) a generative AI assistant powered by speech recognition and large language models, and (3) an optional embodied avatar capable of delivering multimodal feedback through voice and animation. This integrated setup enabled users to engage in conversational learning while immersed in a virtual classroom.

To evaluate the system, eighteen participants experienced three learning conditions: traditional desktop video learning (C1), VR with a text-based AI panel (C2), and VR with a speaking embodied avatar (C3). The results showed that the VR conditions (C2 and C3) significantly enhanced immersion and subjective enjoyment compared to the desktop condition (C1), supporting prior findings on the motivational benefits of immersive media. Participants also tended to prefer the avatar-assisted condition, suggesting a potential affective advantage of embodied social presence.

Despite these subjective benefits, the study did not find significant differences in quiz performance, perceived cognitive load, or reported effort across the three conditions. One possible explanation is that the exposure duration (approximately three minutes per video) may have been too brief to capture measurable learning gains. Additionally, the video-centric format may have limited opportunities for active learning and deeper engagement with the AI assistant. Nonetheless, a modest increase in quiz scores was observed in the AI-assisted conditions, which may suggest small benefits in conceptual reinforcement or attentional support.

Open-ended feedback revealed several challenges. Participants noted that the AI assistant sometimes produced delayed or repetitive responses, and that the avatar's speech occasionally lacked natural emotional tone or contextual alignment. While some users described the avatar as mechanical or uncanny, others found its presence comforting or motivating. These mixed reactions underscore the complexity of designing educational agents that feel human-like while maintaining clarity and responsiveness.

Language accessibility was also identified as a barrier. Although the system supported English dialogue, several participants emphasized the need for multilingual interaction, particularly for non-native speakers. In addition, users suggested that more interactive features—such as in-VR note-taking, guidance prompts, or adaptive quizzes—could enhance the learning experience and support greater learner autonomy.

In sum, this research provides early empirical evidence that embodied generative AI can enhance the affective dimensions of learning in VR. While its cognitive bene-

fits remain inconclusive, the findings suggest that avatar-based interaction may increase motivation and user satisfaction—both of which are essential for sustaining long-term engagement. Future work should focus on improving system responsiveness, speech naturalness, and dialogue personalization. Expanding the system's pedagogical functionalities—such as scaffolding strategies, concept checking, or multimodal input (e.g., diagrams, handwriting)—will also be important for delivering more effective educational support.

This work contributes to the growing field of AI-enhanced immersive education by demonstrating a practical framework for integrating conversational AI, avatars, and video-based instruction into a cohesive VR learning environment. It also points toward new directions for designing emotionally supportive, socially engaging, and technically robust systems for self-paced education in complex domains such as calculus.

References

[1] Y.-S. Su, H.-W. Cheng, and C.-F. Lai, "Study of Virtual Reality Immersive Technology Enhanced Mathematics Geometry Learning," *Frontiers in Psychology*, vol. 13, feb 17 2022. [Online]. Available: http://dx.doi.org/10.3389/fpsyg.2022.760418

[2] M. Simonetti, D. Perri, N. Amato, and O. Gervasi, *Teaching Math with the Help of Virtual Reality*. Springer International Publishing, 2020, pp. 799–809. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-58820-5_57

[3] K. H. Cheong, J. S. Chen, K. Kang, and D. J. Yeo, "Supporting Students' Visualization of Multivariable Calculus Partial Derivatives via Virtual Reality," *Mathematics*, vol. 11, no. 4, p. 831, feb 6 2023. [Online]. Available: http://dx.doi.org/10.3390/math11040831

[4] R. Villena-Taranilla, S. Tirado-Olivares, R. Cózar-Gutiérrez, and J. A. González-Calero, "Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis," *Educational Research Review*, vol. 35, p. 100434, 2 2022. [Online]. Available: http://dx.doi.org/10.1016/j.edurev.2022.100434

[5] D. Hamilton, J. McKechnie, E. Edgerton, and C. Wilson, "Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design," *Journal of Computers in Education*, vol. 8, no. 1, pp. 1–32, jul 11 2020. [Online]. Available: http://dx.doi.org/10.1007/s40692-020-00169-2

[6] K. Kang, S. Kushnarev, W. Wei Pin, O. Ortiz, and J. Chen Shihang, "Impact of Virtual Reality on the Visualization of Partial Derivatives in a Multivariable Calculus Class," *IEEE Access*, vol. 8, pp. 58 940–58 947, 2020. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2020.2982972

[7] R. Sajja, Y. Sermet, M. Cikmaz, D. Cwiertny, and I. Demir, "Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education," *Information*, 2023. [Online]. Available: https://arxiv.org/abs/2309.10892

[8] I. Pesovski, R. Santos, R. Henriques, and V. Trajkovik, "Generative AI for Customizable Learning Experiences," *Sustainability*, vol. 16, no. 7, p. 3034, apr 5 2024. [Online]. Available: http://dx.doi.org/10.3390/su16073034

[9] Z. Bahroun, C. Anane, V. Ahmed, and A. Zacca, "Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational

Settings through Bibliometric and Content Analysis," *Sustainability*, vol. 15, no. 17, p. 12983, aug 29 2023. [Online]. Available: http://dx.doi.org/10.3390/su151712983

[10] H. Qawaqneh, F. B. Ahmad, and A. R. Alawamreh, "The Impact of Artificial Intelligence-Based Virtual Laboratories on Developing Students' Motivation Towards Learning Mathematics," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 18, no. 14, pp. 105–121, jul 31 2023. [Online]. Available: http://dx.doi.org/10.3991/ijet.v18i14.39873

[11] H. Ji, I. Han, and Y. Ko, "A systematic review of conversational AI in language education: focusing on the collaboration with human teachers," *Journal of Research on Technology in Education*, vol. 55, no. 1, pp. 48–63, nov 30 2022. [Online]. Available: http://dx.doi.org/10.1080/15391523.2022.2142873

[12] L. Chen, P. Chen, and Z. Lin, "Artificial Intelligence in Education: A Review," *IEEE Access*, vol. 8, pp. 75 264–75 278, 2020. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2020.2988510

[13] R. O. Davis, T. Park, and J. Vincent, "A Meta-Analytic Review on Embodied Pedagogical Agent Design and Testing Formats," *Journal of Educational Computing Research*, vol. 61, no. 1, pp. 30–67, may 23 2022. [Online]. Available: http://dx.doi.org/10.1177/07356331221100556

[14] N. L. Schroeder and S. D. Craig, "Learning with virtual humans: Introduction to the special issue," *Journal of Research on Technology in Education*, vol. 53, no. 1, pp. 1–7, jan 2 2021. [Online]. Available: http://dx.doi.org/10.1080/15391523.2020.1863114

[15] T. Mizuho, S. Takenaka, T. Narumi, and H. Kuzuoka, "Multiple Self-Avatar Effect: Effects of Using Diverse Self-Avatars on Memory Acquisition and Retention of Sign-Language Gestures," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–15, 2024. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2024.3433498

[16] J. C. Castro-Alonso, P. Ayres, S. Zhang, B. B. de Koning, and F. Paas, "Research Avenues Supporting Embodied Cognition in Learning and Instruction," *Educational Psychology Review*, vol. 36, no. 1, jan 19 2024. [Online]. Available: http://dx.doi.org/10.1007/s10648-024-09847-4

[17] D. Parmar, L. Lin, N. DSouza, S. Jörg, A. E. Leonard, S. B. Daily, and S. V. Babu, "How Immersion and Self-Avatars in VR Affect Learning Programming and Computational Thinking in Middle School Education," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 8, pp. 3698–3713, aug 1 2023. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2022.3169426

[18] Y. O. Cakmak, B. K. Daniel, N. Hammer, O. Yilmaz, E. C. Irmak, and P. Khwaounjoo, "The Human Muscular Arm Avatar as an Interactive Visualization Tool in Learning Anatomy: Medical Students' Perspectives," *IEEE Transactions on Learning Technologies*, vol. 13, no. 3, pp. 593–603, jul 1 2020. [Online]. Available: http://dx.doi.org/10.1109/TLT.2020.2995163

[19] S. Haar, G. Sundar, and A. A. Faisal, "Embodied virtual reality for the study of real-world motor learning," *PLOS ONE*, vol. 16, no. 1, p. e0245717, jan 27 2021. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0245717

[20] I. S. Fitton, D. J. Finnegan, and M. J. Proulx, "Immersive virtual environments and embodied agents for e-learning applications," *PeerJ Computer Science*, vol. 6, p. e315, nov 16 2020. [Online]. Available: http://dx.doi.org/10.7717/peerj-cs.315

[21] S. Chiquet, C. S. Martarelli, D. Weibel, and F. W. Mast, "Learning by teaching in immersive virtual reality – Absorption tendency increases learning outcomes," *Learning and Instruction*, vol. 84, p. 101716, 4 2023. [Online]. Available: http://dx.doi.org/10.1016/j.learninstruc.2022.101716

[22] A. Genay, A. Lecuyer, and M. Hachet, "Being an Avatar "for Real": A Survey on Virtual Embodiment in Augmented Reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 12, pp. 5071–5090, dec 1 2022. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2021.3099290

[23] I. Drori, S. Zhang, R. Shuttleworth, L. Tang, A. Lu, E. Ke, K. Liu, L. Chen, S. Tran, N. Cheng, R. Wang, N. Singh, T. L. Patti, J. Lynch, A. Shporer, N. Verma, E. Wu, and G. Strang, "A neural network solves, explains, and generates university math problems by program synthesis and few-shot learning at human level," *Proceedings of the National Academy of Sciences*, vol. 119, no. 32, aug 2 2022. [Online]. Available: http://dx.doi.org/10.1073/pnas.2123433119

[24] A. Davies, P. Veličković, L. Buesing, S. Blackwell, D. Zheng, N. Tomašev, R. Tanburn, P. Battaglia, C. Blundell, A. Juhász, M. Lackenby, G. Williamson, D. Hassabis, and P. Kohli, "Advancing mathematics by guiding human intuition with AI," *Nature*, vol. 600, no. 7887, pp. 70–74, dec 1 2021. [Online]. Available: http://dx.doi.org/10.1038/s41586-021-04086-x

[25] X. Chen and D. Wu, "Automatic Generation of Multimedia Teaching Materials Based on Generative AI: Taking Tang Poetry as an Example," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 1327–1340, 2024. [Online]. Available: http://dx.doi.org/10.1109/TLT.2024.3378279

[26] M. Menekse, "Envisioning the future of learning and teaching engineering in the artificial intelligence era: Opportunities and challenges," *Journal of Engineering Education*, vol. 112, no. 3, pp. 578–582, jun 20 2023. [Online]. Available: http://dx.doi.org/10.1002/jee.20539

[27] C. K. Boscardin, B. Gin, P. B. Golde, and K. E. Hauer, "Chatgpt and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity," *Academic Medicine*, vol. 99, no. 1, pp. 22–27, aug 31 2023. [Online]. Available: http://dx.doi.org/10.1097/ACM.0000000000005439

[28] S.-C. Kong and Y. Yang, "A Human-Centered Learning and Teaching Framework Using Generative Artificial Intelligence for Self-Regulated Learning Development Through Domain Knowledge Learning in K–12 Settings," *IEEE*

Transactions on Learning Technologies, vol. 17, pp. 1562–1573, 2024. [Online]. Available: http://dx.doi.org/10.1109/TLT.2024.3392830

[29] L. Kohnke, B. L. Moorhouse, and D. Zou, "Chatgpt for Language Teaching and Learning," *RELC Journal*, vol. 54, no. 2, pp. 537–550, apr 3 2023. [Online]. Available: http://dx.doi.org/10.1177/00336882231162868

[30] I. Dergaa, K. Chamari, P. Zmijewski, and H. Ben Saad, "From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing," *Biology of Sport*, vol. 40, no. 2, pp. 615–622, 2023. [Online]. Available: http://dx.doi.org/10.5114/biolsport.2023.125623

[31] U. Mittal, S. Sai, V. Chamola, and D. Sangwan, "A Comprehensive Review on Generative AI for Education," *IEEE Access*, vol. 12, pp. 142 733–142 759, 2024. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2024.3468368

[32] G. Cooper, "Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence," *Journal of Science Education and Technology*, vol. 32, no. 3, pp. 444–452, mar 22 2023. [Online]. Available: http://dx.doi.org/10.1007/s10956-023-10039-y

[33] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, jul 17 2023. [Online]. Available: http://dx.doi.org/10.1186/s41239-023-00411-8

[34] M. Mejeh and M. Rehm, "Taking adaptive learning in educational settings to the next level: leveraging natural language processing for improved personalization," *Educational technology research and development*, vol. 72, no. 3, pp. 1597–1621, feb 21 2024. [Online]. Available: http://dx.doi.org/10.1007/s11423-024-10345-1

[35] R. R. Divekar*, J. Drozdal*, S. Chabot*, Y. Zhou, H. Su, Y. Chen, H. Zhu, J. A. Hendler, and J. Braasch, "Foreign language acquisition via artificial intelligence and extended reality: design and evaluation," *Computer Assisted Language Learning*, vol. 35, no. 9, pp. 2332–2360, feb 15 2021. [Online]. Available: http://dx.doi.org/10.1080/09588221.2021.1879162

[36] T. Xie, R. Liu, Y. Chen, and G. Liu, "Moca: A Motivational Online Conversational Agent for Improving Student Engagement in Collaborative Learning," *IEEE Transactions on Learning Technologies*, vol. 14, no. 5, pp. 653–664, oct 1 2021. [Online]. Available: http://dx.doi.org/10.1109/tlt.2021.3129800

[37] W.-S. Wang, C.-J. Lin, H.-Y. Lee, Y.-M. Huang, and T.-T. Wu, "Integrating feedback mechanisms and ChatGPT for VR-based experiential learning: impacts on reflective thinking and AIoT physical hands-on tasks," *Interactive Learning Environments*, pp. 1–18, jul 15 2024. [Online]. Available: http://dx.doi.org/10.1080/10494820.2024.2375644

[38] M. V. Taylor, Z. Muwaffak, M. R. Penny, B. R. Szulc, S. Brown, A. Merritt, and S. T. Hilton, "Optimising digital twin laboratories with conversational AIs: enhancing immersive training and simulation through virtual reality," *Digital Discovery*, vol. 4, no. 5, pp. 1134–1141, 2025. [Online]. Available: http://dx.doi.org/10.1039/d4dd00330f

[39] S. Y. Liaw, J. Z. Tan, S. Lim, W. Zhou, J. Yap, R. Ratan, S. L. Ooi, S. J. Wong, B. Seah, and W. L. Chua, "Artificial intelligence in virtual reality simulation for interprofessional communication training: Mixed method study," *Nurse Education Today*, vol. 122, p. 105718, 3 2023. [Online]. Available: http://dx.doi.org/10.1016/j.nedt.2023.105718

[40] N. Sardesai, P. Russo, J. Martin, and A. Sardesai, "Utilizing generative conversational artificial intelligence to create simulated patient encounters: a pilot study for anaesthesia training," *Postgraduate Medical Journal*, vol. 100, no. 1182, pp. 237–241, jan 17 2024. [Online]. Available: http://dx.doi.org/10.1093/postmj/qgad137

[41] F. Garzotto, M. Gianotti, A. Patti, F. Pentimalli, and F. Vona, "Empowering Persons with Autism Through Cross-Reality and Conversational Agents," *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 5, pp. 2591–2601, 5 2024. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2024.3372110

[42] R. Hussain, A. Sabir, D.-Y. Lee, S. F. A. Zaidi, A. Pedro, M. S. Abbas, and C. Park, "Conversational AI-based VR system to improve construction safety training of migrant workers," *Automation in Construction*, vol. 160, p. 105315, 4 2024. [Online]. Available: http://dx.doi.org/10.1016/j.autcon.2024.105315

[43] T. C. Luo, A. Aguilera, C. R. Lyles, and C. A. Figueroa, "Promoting Physical Activity Through Conversational Agents: Mixed Methods Systematic Review," *Journal of Medical Internet Research*, vol. 23, no. 9, p. e25486, sep 14 2021. [Online]. Available: http://dx.doi.org/10.2196/25486

[44] A. S. Bergner, C. Hildebrand, and G. Häubl, "Machine Talk: How Verbal Embodiment in Conversational AI Shapes Consumer–Brand Relationships," *Journal of Consumer Research*, vol. 50, no. 4, pp. 742–764, mar 2 2023. [Online]. Available: http://dx.doi.org/10.1093/jcr/ucad014

[45] T. Ribeiro de Oliveira, B. Biancardi Rodrigues, M. Moura da Silva, R. Antonio N. Spinassé, G. Giesen Ludke, M. Ruy Soares Gaudio, G. Iglesias Rocha Gomes, L. Guio Cotini, D. da Silva Vargens, M. Queiroz Schimidt, R. Varejão Andreão, and M. Mestria, "Virtual Reality Solutions Employing Artificial Intelligence Methods: A Systematic Literature Review," *ACM Computing Surveys*, vol. 55, no. 10, pp. 1–29, feb 2 2023. [Online]. Available: http://dx.doi.org/10.1145/3565020

[46] J.-C. Sakdavong, M. Burgues, and N. Huet, "Virtual Reality in Self-regulated Learning: Example in Art Domain," in *Proceedings of the 11th International Conference on Computer Supported Education*. SCITEPRESS

- Science and Technology Publications, 2019, pp. 79–87. [Online]. Available: http://dx.doi.org/10.5220/0007718500790087

[47] Y.-L. Chen and C.-C. Hsu, "Self-regulated mobile game-based English learning in a virtual reality environment," *Computers & Education*, vol. 154, p. 103910, 9 2020. [Online]. Available: http://dx.doi.org/10.1016/j.compedu.2020.103910

[48] Y. Jiang, J. Clarke-Midura, R. S. Baker, L. Paquette, and B. Keller, *How Immersive Virtual Environments Foster Self-Regulated Learning*. IGI Global, 2018, pp. 28–54. [Online]. Available: http://dx.doi.org/10.4018/978-1-5225-3940-7.CH002

[49] M. Sobocinski, D. Dever, M. Wiedbusch, F. Mubarak, R. Azevedo, and S. Järvelä, "Capturing selfregulated learning processes in virtual reality: Causal sequencing of multimodal data," *British Journal of Educational Technology*, vol. 55, no. 4, pp. 1486–1506, sep 30 2023. [Online]. Available: http://dx.doi.org/10.1111/bjet.13393

[50] A. Neo, J. Moura, J. Régis de Araújo, G. Neo, and O. Freitas Júnior, "The Use of Self-Regulation of Learning in Recommender Systems: State-of-the-Art and Research Opportunities," in *Proceedings of the 16th International Conference on Computer Supported Education*. SCITEPRESS - Science and Technology Publications, 2024, pp. 525–532. [Online]. Available: http://dx.doi.org/10.5220/0012619400003693

[51] W. Wang, H. Lee, C. Lin, P. Li, Y. Huang, and T. Wu, "Enhancing students' learning outcomes in selfregulated virtual reality learning environment with learning aid mechanisms," *British Journal of Educational Technology*, vol. 56, no. 1, pp. 366–387, jul 31 2024. [Online]. Available: http://dx.doi.org/10.1111/bjet.13512

[52] W.-L. Wu, Y. Hsu, Q.-F. Yang, and J.-J. Chen, "A Spherical Video-Based Immersive Virtual Reality Learning System to Support Landscape Architecture Students' Learning Performance during the COVID-19 Era," *Land*, vol. 10, no. 6, p. 561, may 27 2021. [Online]. Available: http://dx.doi.org/10.3390/LAND10060561

[53] C.-H. Yang, S.-F. Liu, C.-Y. Lin, and C.-F. Liu, "Immersive Virtual Reality-Based Cardiopulmonary Resuscitation Interactive Learning Support System," *IEEE Access*, vol. 8, pp. 120 870–120 880, 2020. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2020.3006280

[54] S. Koçdar, A. Karadeniz, A. Bozkurt, and K. Buyuk, "Measuring Self-Regulation in Self-Paced Open and Distance Learning Environments," *The International Review of Research in Open and Distributed Learning*, vol. 19, no. 1, 2018. [Online]. Available: https://doi.org/10.19173/irrodl.v19i1.3255

[55] M. Wang, X.-Q. Lyu, Y.-J. Li, and F.-L. Zhang, "Vr content creation and exploration with deep learning: A survey," *Computational Visual Media*, vol. 6, no. 1, pp. 3–28, 3 2020. [Online]. Available: http://dx.doi.org/10.1007/s41095-020-0162-z

[56] B. Marks and J. Thomas, "Adoption of virtual reality technology in higher education: An evaluation of five teaching semesters in a purpose-designed laboratory," *Education and Information Technologies*, vol. 27, no. 1, pp. 1287–1305, jul 8 2021. [Online]. Available: http://dx.doi.org/10.1007/s10639-021-10653-6

[57] A. Gauthier, K. Porayska-Pomsta, I. Dumontheil, S. Mayer, and D. Mareschal, "Manipulating Interface Design Features Affects Children's Stop-And-Think Behaviours in a Counterintuitive-Problem Game," *ACM Transactions on Computer-Human Interaction*, vol. 29, no. 2, pp. 1–21, jan 16 2022. [Online]. Available: http://dx.doi.org/10.1145/3485168

[58] X. Peng, Q. Xu, Y. Chen, C. Zhou, Y. Ge, and N. Li, "An eye tracking study: positive emotional interface design facilitates learning outcomes in multimedia learning?" *International Journal of Educational Technology in Higher Education*, vol. 18, no. 1, jul 20 2021. [Online]. Available: http://dx.doi.org/10.1186/s41239-021-00274-x

[59] A. Pandita and R. Kiran, "The Technology Interface and Student Engagement Are Significant Stimuli in Sustainable Student Satisfaction," *Sustainability*, vol. 15, no. 10, p. 7923, may 12 2023. [Online]. Available: http://dx.doi.org/10.3390/su15107923

[60] S. Alzahrani and L. Roberts, "The effect of visuospatial designing elements of zoomable user interfaces on second language vocabulary acquisition," *System*, p. 102396, 2 2021. [Online]. Available: http://dx.doi.org/10.1016/j.system.2020.102396

[61] J. Chatain, V. Ramp, V. Gashaj, V. Fayolle, M. Kapur, R. W. Sumner, and S. Magnenat, "Grasping Derivatives: Teaching Mathematics through Embodied Interactions using Tablets and Virtual Reality," in *Interaction Design and Children*. ACM, jun 27 2022, pp. 98–108. [Online]. Available: http://dx.doi.org/10.1145/3501712.3529748

[62] R. Wang, R. Lowe, S. Newton, and T. Kocaturk, "Task complexity and learning styles in situated virtual learning environments for construction higher education," *Automation in Construction*, vol. 113, p. 103148, 5 2020. [Online]. Available: http://dx.doi.org/10.1016/j.autcon.2020.103148

[63] E. Scanlon, "Educational Technology Research: Contexts, Complexity and Challenges," *Journal of Interactive Media in Education*, vol. 2021, no. 1, 2021. [Online]. Available: http://dx.doi.org/10.5334/JIME.580

[64] R. F. Kizilcec, J. N. Bailenson, and C. J. Gomez, "The instructor's face in video instruction: Evidence from two large-scale field studies." *Journal of Educational Psychology*, vol. 107, no. 3, pp. 724–739, 8 2015. [Online]. Available: http://dx.doi.org/10.1037/EDU0000013

[65] M. Beege, N. L. Schroeder, S. Heidig, G. D. Rey, and S. Schneider, "The instructor presence effect and its moderators in instructional video: A series of meta-analyses," *Educational Research Review*, vol. 41, p. 100564, 11 2023. [Online]. Available: http://dx.doi.org/10.1016/j.edurev.2023.100564

[66] H. Polat, "Instructors' presence in instructional videos: A systematic review," *Education and Information Technologies*, vol. 28, no. 7, pp. 8537–8569, dec 23 2022. [Online]. Available: http://dx.doi.org/10.1007/s10639-022-11532-4

[67] Y. Y. Ng and A. Przybylek, "Instructor Presence in Video Lectures: Preliminary Findings From an Online Experiment," *IEEE Access*, vol. 9, pp. 36 485–36 499, 2021. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2021.3058735

[68] Z. Yu, "The effect of teacher presence in videos on intrinsic cognitive loads and academic achievements," *Innovations in Education and Teaching International*, vol. 59, no. 5, pp. 574–585, feb 15 2021. [Online]. Available: http://dx.doi.org/10.1080/14703297.2021.1889394

[69] J. C. Castro-Alonso, B. B. de Koning, L. Fiorella, and F. Paas, "Five Strategies for Optimizing Instructional Materials: Instructor- and Learner-Managed Cognitive Load," *Educational Psychology Review*, vol. 33, no. 4, pp. 1379–1407, mar 9 2021. [Online]. Available: http://dx.doi.org/10.1007/s10648-021-09606-9

[70] N. Wenk, J. Penalver-Andres, K. A. Buetler, T. Nef, R. M. Müri, and L. Marchal-Crespo, "Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment," *Virtual Reality*, vol. 27, no. 1, pp. 307–331, aug 16 2021. [Online]. Available: http://dx.doi.org/10.1007/s10055-021-00565-8

[71] Y. Wang and D. Stein, "Effects of online teaching presence on students' cognitive conflict and engagement," *Distance Education*, vol. 42, no. 4, pp. 547–566, oct 2 2021. [Online]. Available: http://dx.doi.org/10.1080/01587919.2021.1987837

[72] N. J. Ahuja, M. Thapliyal, A. Bisht, T. Stephan, R. Kannan, M. S. Al-Rakhami, and M. Mahmud, "An Investigative Study on the Effects of Pedagogical Agents on Intrinsic, Extraneous and Germane Cognitive Load: Experimental Findings With Dyscalculia and Non-Dyscalculia Learners," *IEEE Access*, vol. 10, pp. 3904–3922, 2022. [Online]. Available: http://dx.doi.org/10.1109/access.2021.3115409

[73] O. Nabavinejad and M. Bouchard, "Variant: Limits," 2023, accessed: 2025-06-21. [Online]. Available: https://triseum.com/variant-limits/

[74] P. Gamer, "Modder wires chatgpt into skyrim vr so npcs can roleplay and remember past conversations," 2023, accessed: 2025-06-21, [Online]. Available: https://www.pcgamer.com/modder-wires-chatgpt-into-skyrim-vr-so-npcs-can-roleplay-and-remember-past-conversations/.

[75] I. AI, "Inworld – character engine for ai npcs," 2023, accessed: 2025-06-21. [Online]. Available: https://www.inworld.ai/

[76] C. Inc., "Convai – create conversational ai characters," 2024, accessed: 2025-06-21. [Online]. Available: https://www.convai.com/

[77] EleVR, "Calcflow – a vr vector calculus experience," 2016, accessed: 2025-06-21. [Online]. Available: https://github.com/EleVR/calcflow

[78] M. Insight, "Calcplot3d – 3d graphing tool for multivariable calculus," 2020, accessed: 2025-06-21. [Online]. Available: https://www.math3d.org/

[79] Unknown, "Calcvr – vr tools for teaching mathematics," 2023, replace with accurate source or academic paper. [Online]. Available: https://calcvr.org/

[80] "Spatial analytics for vr/ar applications," https://cognitive3d.com, 2024, retrieved January 14, 2025.

[81] H. S. Lai, "Unity system repository," 2025, accessed: 2025-06-21. [Online]. Available: https://github.com/huiishan99/uoa-master-research-unity

[82] I. Meta Platforms, "Meta quest link," 2025, accessed: 2025-06-21. [Online]. Available: https://www.meta.com/en-gb/help/quest/1517439565442928

[83] Khan Academy, "Differential calculus – unit 2: Derivatives: Definition and basic rules," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro

[84] M. Corporation, "Clipchamp video editor," 2025, accessed: 2025-06-21. [Online]. Available: https://clipchamp.com

[85] Khan Academy, "Secant lines and average rate of change," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro/dc-diff-calc-intro/v/secant-lines-and-average-rate-of-change

[86] ——, "Derivative as slope of a curve," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro/dc-diff-calc-intro/v/derivative-as-slope-of-curve

[87] ——, "The derivative and tangent line equations," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro/dc-diff-calc-intro/v/derivative-as-slope-of-tangent-line

[88] U. Technologies, "Video player component documentation," 2025, accessed: 2025-06-21. [Online]. Available: https://docs.unity3d.com/2022.3/Documentation/Manual/class-VideoPlayer.html

[89] Khan Academy, "Secants and average rate of change (exercise)," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro/dc-diff-calc-intro/e/secants-and-average-rate-of-change

[90] ——, "Graphs of functions and their derivatives (exercise)," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro/dc-diff-calc-intro/e/graphs-of-functions-and-their-derivatives

[91] ——, "Derivative at a point as slope of tangent line (exercise)," 2024, accessed: 2025-06-17. [Online]. Available: https://www.khanacademy.org/math/differential-calculus/dc-diff-intro/dc-diff-calc-intro/e/derivative-at-a-point-as-slope-of-tangent-line

[92] OpenAI, “Openai assistants api playground,” 2024, accessed: 2025-06-17. [Online]. Available: https://platform.openai.com/playground/assistants

[93] P. Project, “Flask – python web framework,” 2025, accessed: 2025-06-21. [Online]. Available: https://flask.palletsprojects.com/

[94] U. Technologies, “Unitywebrequest api reference,” 2025, accessed: 2025-06-21. [Online]. Available: https://docs.unity3d.com/ScriptReference/Networking.UnityWebRequest.html

[95] OpenAI, “Openai python sdk,” 2025, accessed: 2025-06-21. [Online]. Available: https://github.com/openai/openai-python

[96] U. Technologies, “Unity microphone api reference,” 2025, accessed: 2025-06-21. [Online]. Available: https://docs.unity3d.com/ScriptReference/Microphone.html

[97] Macoron, “whisper.unity – unity integration of whisper,” 2025, accessed: 2025-06-21. [Online]. Available: https://github.com/Macoron/whisper.unity

[98] ggerganov, “Whisper base.en model on hugging face,” 2024, accessed: 2025-06-17. [Online]. Available: https://huggingface.co/ggerganov/whisper.cpp

[99] U. Technologies, “Textmeshpro documentation,” 2025, accessed: 2025-06-21. [Online]. Available: https://docs.unity3d.com/Packages/com.unity.textmeshpro@4.0/manual/index.html

[100] A. W. Services, “Amazon polly documentation,” 2025, accessed: 2025-06-21. [Online]. Available: https://docs.aws.amazon.com/polly/

[101] I. Meta Platforms, “Oculus lipsync sdk for unity,” 2025, accessed: 2025-06-21. [Online]. Available: https://developer.oculus.com/documentation/unity/audio-ovrlipsync-unity/

[102] Wolf3D, “Ready player me – avatar creator,” 2025, accessed: 2025-06-21. [Online]. Available: https://readyplayer.me/

[103] Adobe, “Mixamo – 3d animation platform,” 2025, accessed: 2025-06-21. [Online]. Available: https://www.mixamo.com/

[104] R. P. Me, “Ready player me animation library,” https://github.com/readyplayerme/animation-library/tree/master, 2023, accessed: 2025-06-21.

[105] U. Technologies, *Unity Manual: Animation System*, https://docs.unity3d.com/2022.3/Documentation/Manual/AnimationSection.html, Unity, 2023, accessed: 2025-06-21.

[106] H. Regenbrecht and T. Schubert, “Real and Illusory Interactions Enhance Presence in Virtual Environments,” *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 4, pp. 425–434, 8 2002. [Online]. Available: http://dx.doi.org/10.1162/105474602760204318

[107] S. G. Hart, “Nasa-Task Load Index (NASA-TLX); 20 Years Later,” *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, 10 2006. [Online]. Available: http://dx.doi.org/10.1177/154193120605000909

[108] M. Hassenzahl, M. Burmester, and F. Koller, *AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität*. Vieweg+Teubner Verlag, 2003, pp. 187–196. [Online]. Available: http://dx.doi.org/10.1007/978-3-322-80058-9_19

Chapter 7

Appendices

This appendix section contains supplementary materials supporting the main findings of this thesis. It includes:

- **Appendix A: Log Data** – Raw interaction logs between users and the AI tutor, categorized by experimental conditions.
- **Appendix B: Questionnaire Survey** – Full survey form used for evaluating participant demographics, immersion, cognitive load, and user preferences.
- **Appendix C: Quiz for Research** – Post-video quizzes designed to assess learners' comprehension of calculus concepts across different conditions.
- **Appendix D: Participants Feedback** – Feedback from all participants on immersion, usability, avatar realism, and improvement suggestions.

All documents have been scaled down for layout consistency.

7.1 Appendix A: Log Data

Participant ID	Condition 2	Condition 3
P1	[2025-05-27 15:03:38] AvatarToggle: Hide [2025-05-27 15:04:23] Question: Okay. Ah, hello. Okay. [2025-05-27 15:04:23] Response: Hi there! How can I help you today? [2025-05-27 15:04:23] ResponseTime: 5.06s	[2025-05-27 15:10:44] AvatarToggle: Active [2025-05-27 15:11:07] Question: Okay. Hello. [2025-05-27 15:11:07] Response: Hello again! What would you like to talk about? [2025-05-27 15:11:07] ResponseTime: 4.88s [2025-05-27 15:16:40] Question: Oh. Okay. Okay. Okay. Okay. [2025-05-27 15:16:40] Response: It seems like you're thinking. Take your time! If you have a question or topic in mind, just let me know. [2025-05-27 15:16:40] ResponseTime: 6.46s
P2	[2025-05-28 09:30:04] AvatarToggle: Hide [2025-05-28 09:31:50] Question: Hello. Why? What is tangent like? [2025-05-28 09:31:50] Response: A tangent line is a straight line that touches a curve at a single point. It represents the slope of the curve at that point, showing how the function is changing there. In terms of derivatives, the derivative at a point gives you the slope of the tangent line at that point. [2025-05-28 09:31:50] ResponseTime: 5.17s	[2025-05-28 08:43:51] Initial AvatarToggle: Active [2025-05-28 09:21:16] Question: Thanks for the good one. I just want to just say okay. Hello, hello boys. Hey, how are you? I'm not going to be a teacher. I'm not going to be a teacher. [inaudible] [2025-05-28 09:21:16] Response: It sounds like you're just having a casual chat! If you have any math questions or need help with derivatives, feel free to ask. I'm here to help! [2025-05-28 09:21:16] ResponseTime: 5.53s [2025-05-28 09:25:51] Question: Hello, do you know what is data X? That means the change of X not date. [2025-05-28 09:25:51] Response: Yes, "data X" usually refers to the change in the variable X, often represented as ΔX (delta X). In calculus, this concept relates to derivatives, which measure how a function changes as its input changes. If you have more specific questions about it, let me know! [2025-05-28 09:25:51] ResponseTime: 6.21s
P3	[2025-05-28 10:53:14] AvatarToggle: Hide	[2025-05-28 10:31:17] AvatarToggle: Active [2025-05-28 10:31:17] Initial AvatarToggle: Active [2025-05-28 10:38:22] Question: (speaking in foreign language) (speaking in foreign language) I can speak English. (speaking in foreign language) [2025-05-28 10:38:22] Response: That's great! How can I help you with math today? [2025-05-28 10:38:22] ResponseTime: 6.65s [2025-05-28 10:53:03] AvatarToggle: Active [2025-05-28 10:53:03] Initial AvatarToggle: Active
P4	[2025-05-28 13:07:54] AvatarToggle: Hide [2025-05-28 13:08:17] Question: [BLANK_AUDIO] [BLANK_AUDIO] Light high control circuit. [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [2025-05-28 13:08:17] Response: It looks like there might be some confusion. If you have a specific math question or topic you'd like to discuss, please type it out! [2025-05-28 13:08:17] ResponseTime: 5.53s	[2025-05-28 12:42:20] AvatarToggle: Active [2025-05-28 12:42:20] Initial AvatarToggle: Active [2025-05-28 13:00:51] Question: [BLANK_AUDIO] [BLANK_AUDIO] Okay this. [BLANK_AUDIO] [2025-05-28 13:00:51] Response: It seems like there might be some audio issues. If you have a math question, feel free to type it out! [2025-05-28 13:00:51] ResponseTime: 7.81s [2025-05-28 13:05:00] Question: [BLANK_AUDIO] [BLANK_AUDIO] [2025-05-28 13:05:00] Response: I can't hear anything right now. If you have a question about derivatives or any math topic, just let me know! [2025-05-28 13:05:00] ResponseTime: 4.45s

P5	[2025-05-28 14:50:02] AvatarToggle: Active [2025-05-28 14:50:02] Initial AvatarToggle: Active [2025-05-28 14:50:12] AvatarToggle: Hide [2025-05-28 15:03:37] Question: Okay, here's what you want to go. I'll tell you what, pay it a go. I don't know, cancer pain. Okay. [2025-05-28 15:03:37] Response: It seems like your message is a bit unclear. Could you please clarify what you need help with regarding derivatives? I'm here to assist! [2025-05-28 15:03:37] ResponseTime: 6.48s [2025-05-28 15:03:53] Question: Okay, here's what you want to go. I'll tell you what, pay it a go. I don't know, cancer pain. Okay. [2025-05-28 15:03:53] Response: I'm not quite sure what you're asking. If you have a specific question about derivatives or a math problem, feel free to share! [2025-05-28 15:03:53] ResponseTime: 5.46s	[2025-05-28 15:10:35] AvatarToggle: Active [2025-05-28 15:11:23] Question: Okay, here's what you want to go. I'll tell you what, pay it a go. I don't know, cancer pain. Okay. [2025-05-28 15:11:23] Response: It seems like there might be some confusion. If you have a question about derivatives or need help with a math topic, just let me know! I'm here to help. [2025-05-28 15:11:23] ResponseTime: 6.11s
P6	[2025-05-29 12:56:02] AvatarToggle: Active [2025-05-29 12:56:02] Initial AvatarToggle: Active [2025-05-29 12:56:05] AvatarToggle: Hide [2025-05-29 13:01:39] Question: Hello. Thank you. [inaudible] [2025-05-29 13:01:39] Response: Hi there! How can I help you today? If you have any math questions, feel free to ask! [2025-05-29 13:01:39] ResponseTime: 5.99s	[2025-05-29 13:13:32] AvatarToggle: Active [2025-05-29 13:14:02] Question: Hello. [2025-05-29 13:14:02] Response: Hello again! How can I assist you today? [2025-05-29 13:14:02] ResponseTime: 7.98s
P7	[2025-05-29 20:43:24] AvatarToggle: Active [2025-05-29 20:43:24] Initial AvatarToggle: Active [2025-05-29 20:47:21] AvatarToggle: Hide [2025-05-29 21:00:08] Question: Hello? Oh yes, I see that. Thank you. [2025-05-29 21:00:08] Response: You're welcome! How can I assist you today? [2025-05-29 21:00:08] ResponseTime: 6.58s [2025-05-29 21:06:04] Question: Could you please tell me what is teammate me? [2025-05-29 21:06:04] Response: It seems like you might be asking about a specific term or concept. Could you clarify what you mean by "teammate me"? Are you looking for information on teamwork or collaboration? [2025-05-29 21:06:04] ResponseTime: 7.28s [2025-05-29 21:07:30] Question: [BLANK_AUDIO] Do you know the question? The answer in question two, do you know that? [BLANK_AUDIO] [2025-05-29 21:07:30] Response: I can't hear the audio, but if you provide me with the question or context, I can help you find the answer! [2025-05-29 21:07:30] ResponseTime: 5.60s	[2025-05-29 21:12:31] AvatarToggle: Active [2025-05-29 21:12:57] Question: How are you? What's your name? [2025-05-29 21:12:57] Response: I'm doing well, thank you! I don't have a name, but you can call me your math tutor. How can I help you today? [2025-05-29 21:12:57] ResponseTime: 7.78s [2025-05-29 21:18:40] Question: What is tangent line? [2025-05-29 21:18:40] Response: A tangent line is a straight line that touches a curve at a single point. It represents the slope of the curve at that point, showing how the curve is changing there. [2025-05-29 21:18:40] ResponseTime: 4.87s [2025-05-29 21:21:00] Question: What is the graph of function H? [2025-05-29 21:21:00] Response: I need more details about function H to help you. Could you describe the function or provide its equation? [2025-05-29 21:21:00] ResponseTime: 4.99s [2025-05-29 21:22:11] Question: The function H passes through the point. [2025-05-29 21:22:11] Response: Could you please provide the specific point or more details about function H? That way, I can help you better! [2025-05-29 21:22:11] ResponseTime: 6.20s

P8	[2025-05-30 11:20:22] AvatarToggle: Hide [2025-05-30 11:24:11] Question: Hey, I want to know what the mean is seeing the estimate. I want to know what the mean is. I want to know what the mean is. [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] Reduce the clear. [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [BLANK_AUDIO] [2025-05-30 11:24:11] Response: It seems like there might be some confusion. The mean is simply the average of a set of numbers. If you have specific numbers or a context in mind, feel free to share, and I can help you calculate or understand it better! [2025-05-30 11:24:11] ResponseTime: 4.78s [2025-05-30 11:25:13] Question: Hello, how are you? [BLANK_AUDIO] [2025-05-30 11:25:13] Response: I'm doing well, thank you! How about you? How can I assist you today? [2025-05-30 11:25:13] ResponseTime: 9.63s [2025-05-30 11:30:40] Question: Hi, I want to ask you for the question, the tangent line to the graph of function h and the point minus 2, minus 4. passes through the point, 5, find h minus 2. [2025-05-30 11:30:40] Response: To find h(-2), you need the slope of the tangent line at that point. If you have the slope or the equation of the tangent line, you can use it to find h(-2). Do you have that information? [2025-05-30 11:30:40] ResponseTime: 5.88s [2025-05-30 11:31:19] Question: I want to know this. [2025-05-30 11:31:19] Response: To find h(-2), we need the slope of the tangent line at that point. If you can provide the slope or the equation of the tangent line, I can help you find h(-2). [2025-05-30 11:31:19] ResponseTime: 5.97s	[2025-05-30 10:51:18] AvatarToggle: Active [2025-05-30 10:51:18] Initial AvatarToggle: Active [2025-05-30 11:10:38] Question: Hey, hello. [BLANK_AUDIO] [2025-05-30 11:10:38] Response: Hi there! How can I help you today? [2025-05-30 11:10:38] ResponseTime: 5.34s [2025-05-30 11:15:21] Question: Hey, I want to know what's about to get. This is the X mismatch. [2025-05-30 11:15:21] Response: Could you clarify what you mean by "X mismatch"? Are you referring to a specific math problem or concept related to derivatives? [2025-05-30 11:15:21] ResponseTime: 5.33s [2025-05-30 11:16:17] Question: Sorry, can you tell me your name? [2025-05-30 11:16:17] Response: I'm your friendly math tutor! You can just call me Tutor. How can I assist you with math today? [2025-05-30 11:16:17] ResponseTime: 5.53s [2025-05-30 11:17:04] Question: Hey, I want to know what the mean is... [2025-05-30 11:17:04] Response: The mean is the average of a set of numbers. You find it by adding all the numbers together and then dividing by how many numbers there are. Are you looking for something specific about the mean? [2025-05-30 11:17:04] ResponseTime: 8.36s

P9	[2025-05-31 15:18:14] AvatarToggle: Hide [2025-05-31 15:18:46] Question: Are you recording? Oh yeah, I think it's recording now. Yes. [2025-05-31 15:18:46] Response: It sounds like you're checking on a recording. If you have any questions or need help with math, just let me know! [2025-05-31 15:18:46] ResponseTime: 3.94s [2025-05-31 15:22:38] Question: The tangent line to the graph of function H. At point minus 2, minus 4 passes through the point. 1, 5, find H of minus 2. [BLANK_AUDIO] [2025-05-31 15:22:38] Response: Error code: 400 - {'error': {'message': "Can't add messages to thread_nteymBbvs2QictA1hoTZKTAm while a run run_IyrFNylwFhdiDJKdWdwipbjk is active.", 'type': 'invalid_request_error', 'param': None, 'code': None}} [2025-05-31 15:22:38] ResponseTime: 0.97s [2025-05-31 15:22:40] Question: The tangent line to the graph of function H. At point minus 2, minus 4 passes through the point. 1, 5, find H of minus 2. [BLANK_AUDIO] [2025-05-31 15:22:40] Response: To find H(-2), you need the slope of the tangent line at that point. The slope can be determined using the two points: (-2, -4) and (1, 5). Calculate the slope as (5 - (-4)) / (1 - (-2)) = 9 / 3 = 3. Now, use the point-slope form of the line: y - y1 = m(x - x1). Plugging in the point (-2, -4) and the slope 3 gives you the equation of the tangent line. Since H(-2) is the y-value at that point, it is -4. [2025-05-31 15:22:40] ResponseTime: 3.43s	[2025-05-31 15:03:55] AvatarToggle: Active [2025-05-31 15:03:55] Initial AvatarToggle: Active [2025-05-31 15:05:06] Question: Yeah, it's recording, I think. And I'll stop it now. [2025-05-31 15:05:06] Response: It sounds like you're working on something interesting! If you have any questions about derivatives or need help with math, feel free to ask! [2025-05-31 15:05:06] ResponseTime: 11.96s [2025-05-31 15:05:31] AvatarToggle: Hide [2025-05-31 15:05:32] AvatarToggle: Active [2025-05-31 15:09:11] Question: What is the average rate of change of g of x equal to 2x over the interval? [2025-05-31 15:09:11] Response: To find the average rate of change of g(x) = 2x over an interval, you need the values of g at the endpoints of that interval. The average rate of change is calculated by taking the difference in the function values and dividing by the difference in the x-values. If you provide the specific interval, I can help you calculate it! [2025-05-31 15:09:11] ResponseTime: 6.35s [2025-05-31 15:09:59] Question: Help me to calculate it. [2025-05-31 15:09:59] Response: Please provide the specific interval for x, like [a, b], so I can help you calculate the average rate of change for g(x) = 2x. [2025-05-31 15:09:59] ResponseTime: 5.32s [2025-05-31 15:10:04] Question: Help me to calculate it. [2025-05-31 15:10:04] Response: I still need the specific interval, like [a, b], to calculate the average rate of change. Could you please provide that? [2025-05-31 15:10:04] ResponseTime: 4.09s [2025-05-31 15:10:27] Question: Thank you. The interval is from 3 to 10. [2025-05-31 15:10:27] Response: Great! For the interval [3, 10], first, calculate g(3) and g(10). g(3) = 2(3) = 6 and g(10) = 2(10) = 20. Now, the average rate of change is (g(10) - g(3)) / (10 - 3) = (20 - 6) / (7) = 14 / 7 = 2. So, the average rate of change is 2. [2025-05-31 15:10:27] ResponseTime: 0.07s [2025-05-31 15:10:31] Question: Thank you. The interval is from 3 to 10. [2025-05-31 15:10:31] Response: I already calculated the average rate of change for the interval [3, 10]. It is 2. If you have any more questions or need further assistance, feel free to ask! [2025-05-31 15:10:31] ResponseTime: 3.86s

		[2025-05-31 15:46:59] AvatarToggle: Active [2025-05-31 15:46:59] Initial AvatarToggle: Active [2025-05-31 15:52:42] Question: you. How are you? [BLANK_AUDIO] [2025-05-31 15:52:42] Response: I'm just a program, but I'm here and ready to help you! If you have any questions or need assistance with math, feel free to ask! [2025-05-31 15:52:42] ResponseTime: 5.42s [2025-05-31 15:57:13] Question: Will you explain the question? [2025-05-31 15:57:13] Response: Sure! The question asks for the value of the function H at the point (-2, -4) and mentions that the tangent line at this point also passes through another point (1, 5). You need to find the slope of the tangent line using these two points and then confirm that H(-2) is -4, which is given in the question. If you need more details, just let me know! [2025-05-31 15:57:13] ResponseTime: 8.65s [2025-05-31 15:58:32] Question: Uh quiz one, what is the outrage rate of? Econch-o-fo-x. Over the internal one. 1, 10, 3, and... [2025-05-31 15:58:32] Response: It seems like there might be a typo in your question. If you're asking for the average rate of change of a function over the interval [1, 3], please provide the correct function, and I can help you calculate it! [2025-05-31 15:58:32] ResponseTime: 4.98s [2025-05-31 16:00:00] Question: Function is... uh... to... much... it's called... X-Co. G...
	[2025-05-31 16:05:09] AvatarToggle: Hide [2025-05-31 16:05:48] Question: How are you? [2025-05-31 16:05:48] Response: I'm here and ready to help you! How can I assist you today? [2025-05-31 16:05:48] ResponseTime: 7.84s [2025-05-31 16:10:52] Question: I have a question. As a teammate, H prime 10. You [2025-05-31 16:10:52] Response: It seems like you're asking about the derivative of a function H at x = 10, denoted as H'(10). If you provide the function H, I can help you find its derivative at that point! [2025-05-31 16:10:52] ResponseTime: 6.78s [2025-05-31 16:11:56] Question: have a question. Why equal hx? [2025-05-31 16:11:56] Response: It looks like you're asking about a function where H equals some expression involving x. If you clarify what H is equal to, I can help explain or find the derivative for you! [2025-05-31 16:11:56] ResponseTime: 6.53s	[2025-05-31 16:00:00] Response: It sounds like you're trying to describe a function, but I'm not quite clear on the name. If you can clarify the function you're referring to, I can help you find the average rate of change over the interval you mentioned! [2025-05-31 16:00:00] ResponseTime: 6.76s [2025-05-31 16:00:40] Question: uh... the function is... two x... [2025-05-31 16:00:40] Response: Got it! If the function is g(x) = 2x, and you want to find the average rate of change over the interval [1, 3], here's how to do it: First, calculate g(1) = 2(1) = 2 and g(3) = 2(3) = 6. Now, the average rate of change is (g(3) - g(1)) / (3 - 1) = (6 - 2) / (2) = 4 / 2 = 2. So, the average rate of change is 2. [2025-05-31 16:00:40] ResponseTime: 7.28s
P10		

P11	[2025-06-02 12:56:55] AvatarToggle: Active [2025-06-02 12:56:55] Initial AvatarToggle: Active [2025-06-02 12:56:59] AvatarToggle: Hide [2025-06-02 13:02:33] Question: How are you? Make a pleasure. Take a bite. [2025-06-02 13:02:33] Response: I'm just a math tutor, but I'm here to help you with any math questions you have! What do you need assistance with today? [2025-06-02 13:02:33] ResponseTime: 6.38s	[2025-06-02 13:08:58] AvatarToggle: Active [2025-06-02 13:09:26] Question: Hey girl, what's up? [2025-06-02 13:09:26] Response: I'm here to help you with math! Do you have any questions about derivatives or anything else? [2025-06-02 13:09:26] ResponseTime: 4.96s [2025-06-02 13:14:13] Question: Hey girl, could you tell me how to estimate the change rate of the function at some point? [2025-06-02 13:14:13] Response: To estimate the change rate of a function at a point, you can find the derivative at that point. The derivative gives you the slope of the tangent line, which represents the rate of change. Do you have a specific function in mind? [2025-06-02 13:14:13] ResponseTime: 5.19s [2025-06-02 13:14:58] Question: Hey girl, could you tell me how to estimate the change rate of the function at some point? Do you know? Yeah, think about a function that y equal to 2 times x. [2025-06-02 13:14:58] Response: For the function y = 2x, the derivative is simply 2. This means the change rate is constant at 2, regardless of the value of x. So, at any point, the function is increasing at a rate of 2 units of y for every 1 unit of x. [2025-06-02 13:14:58] ResponseTime: 6.00s [2025-06-02 13:15:30] Question: Hey girl, could you tell me how to estimate the change rate of the function at some point? Do you know? Yeah, think about a function that y equal to 2 times x. Okay, thank you. Okay. [2025-06-02 13:15:30] Response: You're welcome! If you have any more questions about derivatives or anything else, feel free to ask. I'm here to help! [2025-06-02 13:15:30] ResponseTime: 5.63s
P12	[2025-06-02 13:31:43] AvatarToggle: Active [2025-06-02 13:31:43] Initial AvatarToggle: Active [2025-06-02 13:32:32] AvatarToggle: Hide [2025-06-02 13:38:01] Question: into code pass to send. Sosamal. Hello? How are you? Hello, you answer. I don't know what to say. [2025-06-02 13:38:01] Response: It seems like you're looking for help with something specific. Could you clarify what you need assistance with? [2025-06-02 13:38:01] ResponseTime: 5.74s	[2025-06-02 13:48:32] AvatarToggle: Active [2025-06-02 13:49:09] Question: Would you tell me about something about this system? about this term? software? [2025-06-02 13:49:09] Response: Could you please specify which term or software you're referring to? That way, I can provide you with the information you need. [2025-06-02 13:49:09] ResponseTime: 5.24s [2025-06-02 13:54:53] Question: The tangent line to the graph is a function h and it passes through the point negative 2 and negative 4. And its tangent line passes through the point 1, 5. Could you find h? h? [2025-06-02 13:54:53] Response: To find the function h, we need to determine the slope of the tangent line at the given points. The slope can be calculated using the two points you provided. Once we have the slope, we can use the point-slope form of a line to find the equation of the tangent line. Would you like to go through the calculations together? [2025-06-02 13:54:53] ResponseTime: 5.82s
P13	[2025-06-03 10:53:10] AvatarToggle: Active [2025-06-03 10:53:10] Initial AvatarToggle: Active [2025-06-03 10:53:51] AvatarToggle: Hide [2025-06-03 11:04:56] Question: Hello [2025-06-03 11:04:56] Response: Error code: 400 - {'error': {'message': "Can't add messages to thread_F4EAY1s9xaEJ3slVbpWPphh1 while a run run_iVuxr6IoiyeseutZopvDLmW3 is active.", 'type': 'invalid_request_error', 'param': None, 'code': None}} [2025-06-03 11:04:56] ResponseTime: 1.01s [2025-06-03 11:04:59] Question: Hello [2025-06-03 11:04:59] Response: Hi there! How can I help you today? [2025-06-03 11:04:59] ResponseTime: 4.58s [2025-06-03 11:05:10] Question: Hello [2025-06-03 11:05:10] Response: Hello again! What would you like to talk about? [2025-06-03 11:05:10] ResponseTime: 8.93s	[2025-06-03 11:10:10] AvatarToggle: Active [2025-06-03 11:10:34] Question: Hi. Toot it. [2025-06-03 11:10:34] Response: I'm here to help! What math topic do you want to discuss? [2025-06-03 11:10:34] ResponseTime: 4.98s [2025-06-03 11:10:40] Question: Hi. To the register. Yeah. Did. Ah. [2025-06-03 11:10:40] Response: It sounds like you might be trying to say something specific. Can you clarify what you need help with? [2025-06-03 11:10:40] ResponseTime: 5.16s [2025-06-03 11:10:53] Question: Hi. To the register. Yeah. Did. Ah. [2025-06-03 11:10:53] Response: I'm not quite sure what you mean. Could you provide a bit more detail or ask a specific question? [2025-06-03 11:10:53] ResponseTime: 4.32s

P14	[2025-06-03 13:16:23] AvatarToggle: Hide [2025-06-03 13:16:48] Question: What's up out there, whether today? [2025-06-03 13:16:48] Response: I can't check the weather, but I can help you with math or any questions you have! What do you need help with? [2025-06-03 13:16:48] ResponseTime: 5.47s	[2025-06-03 12:52:49] AvatarToggle: Active [2025-06-03 12:52:49] Initial AvatarToggle: Active [2025-06-03 13:10:11] Question: Hello, nice to meet you. [2025-06-03 13:10:11] Response: Hi there! Nice to meet you too. How can I help you today? [2025-06-03 13:10:11] ResponseTime: 5.99s
P15	[2025-06-06 15:15:11] AvatarToggle: Hide [2025-06-06 15:15:35] Question: Hi. Can you hear me? I'll press A through X descent. Okay. Okay. So we want to figure out the average slope, the rate of change of Y over the rate of change of X. So if I could double check with you, it would be 20 minus 6. Over 10 minus 3. So I think that's. 14 over 7 or 2. And that makes sense logically. Do you want to double check me? Maybe that's too long. Mm hmm. Oh, take it down. You you you you you you You You You Oh, yeah, no, no. Yeah. It's time to see. Panel. What do you mean? [2025-06-06 15:15:35] Response: I can hear you! It sounds like you're working through the average slope calculation again. Your math is correct: 20 minus 6 over 10 minus 3 is indeed 14 over 7, which equals 2. If you have more questions or need further clarification, just let me know! [2025-06-06 15:15:35] ResponseTime: 6.19s [2025-06-06 15:17:41] Question: Hi. So I am thinking we would just take the change of X. Sorry, the change of Y over the change of X just as we did both. That should be 9 over... [2025-06-06 15:17:41] Response: Oops, something went wrong. Please try again. [2025-06-06 15:17:41] ResponseTime: 15.02s [2025-06-06 15:18:14] Question: Hello. I believe it is 9/3 or 3. Can you check this for me? [2025-06-06 15:18:14] Response: Sure! If you're calculating the change in Y as 9 and the change in X as 3, then 9 divided by 3 is indeed 3. Great job! [2025-06-06 15:18:14] ResponseTime: 5.85s	[2025-06-06 14:52:02] AvatarToggle: Active [2025-06-06 14:52:02] Initial AvatarToggle: Active [2025-06-06 15:05:29] Question: I hit it. Okay. Yeah, I see. Cool. So that's the AI. Okay. I'm going to take it on the other side. [2025-06-06 15:05:29] Response: It sounds like you're exploring something interesting! If you have any math questions, especially about derivatives, feel free to ask. I'm here to help! [2025-06-06 15:05:29] ResponseTime: 6.20s [2025-06-06 15:10:03] Question: Hi. Can you hear me? I'll press A through X descent. Okay. Okay. So we want to figure out the average slope, the rate of change of Y over the rate of change of X. So if I could double check with you, it would be 20 minus 6. Over 10 minus 3. So I think that's. 14 over 7 or 2. And that makes sense. [2025-06-06 15:10:03] Response: Yes, you're on the right track! The average slope, or rate of change, is calculated by taking the difference in Y values divided by the difference in X values. So, 20 minus 6 over 10 minus 3 gives you 14 over 7, which simplifies to 2. Great job! [2025-06-06 15:10:03] ResponseTime: 6.22s

P16	[2025-06-06 16:06:19] AvatarToggle: Hide [2025-06-06 16:06:41] Question: Hi, can you hear me? [BLANK_AUDIO] [2025-06-06 16:06:41] Response: I can hear you! If you have more questions or need help with something, just let me know! [2025-06-06 16:06:41] ResponseTime: 5.90s [2025-06-06 16:10:37] Question: What is the meaning of hash prime? [2025-06-06 16:10:37] Response: In mathematics, "hash prime" isn't a standard term. You might be thinking of "hash" in computer science, which relates to data structures, or "prime," which refers to prime numbers. If you meant something specific, could you clarify? [2025-06-06 16:10:37] ResponseTime: 4.30s [2025-06-06 16:11:17] Question: I'm asking what is the meaning of hash prime s? No, hash prime x. [BLANK_AUDIO] [2025-06-06 16:11:17] Response: It sounds like you're referring to the notation "hash prime x," which often represents the derivative of a function. In calculus, the prime symbol (') indicates that you're looking at the rate of change of a function with respect to its variable. So, hash prime x usually means the derivative of a function with respect to x. [2025-06-06 16:11:17] ResponseTime: 5.60s	[2025-06-06 15:49:54] AvatarToggle: Active [2025-06-06 15:49:54] Initial AvatarToggle: Active [2025-06-06 15:55:37] Question: Hello? Hello? Hello audio? Oh hello, yes. [inaudible] [2025-06-06 15:55:37] Response: Hi there! It seems like you're trying to communicate. How can I help you today? If you have any math questions, especially about derivatives, feel free to ask! [2025-06-06 15:55:37] ResponseTime: 7.20s [2025-06-06 16:00:55] Question: Hello? Can you hear me? Okay, I have a question. What does the question mean? What? What is the question? [laughs] What is the meaning of average rate? [2025-06-06 16:00:55] Response: The average rate refers to how much something changes over a specific interval. In calculus, it's often related to the average rate of change of a function, which is calculated by finding the difference in the function's values at two points and dividing by the difference in the input values. It's like finding the slope of a line connecting two points on a graph! [2025-06-06 16:00:55] ResponseTime: 5.89s [2025-06-06 16:02:17] Question: Hello? Can you hear me? I have a... a quick question? What is the average rate of change of GS equal to 2S over the interval 3 and 10? [2025-06-06 16:02:17] Response: To find the average rate of change of the function GS = 2S over the interval from 3 to 10, you subtract the function's value at 3 from its value at 10, then divide by the difference in the S values. So, calculate GS(10) - GS(3) and divide that by 10 - 3. Would you like me to help with the calculations? [2025-06-06 16:02:17] ResponseTime: 6.26s [2025-06-06 16:03:07] Question: Yes, please tell me. No, I mean, please, me. The calculation. [2025-06-06 16:03:07] Response: The average rate of change of GS = 2S over the interval from 3 to 10 is 2. [2025-06-06 16:03:07] ResponseTime: 7.58s [2025-06-06 16:03:32] Question: Tell me why. [2025-06-06 16:03:32] Response: The average rate of change is essentially the slope of the line connecting the two points on the graph of the function. Since GS = 2S is a linear function, the slope is constant. So, for every unit increase in S, GS increases by 2, which gives us an average rate of change of 2 over any interval. [2025-06-06 16:03:32] ResponseTime: 6.09s
P17	[2025-06-09 08:49:20] AvatarToggle: Active [2025-06-09 08:49:20] Initial AvatarToggle: Active [2025-06-09 08:49:22] AvatarToggle: Hide [2025-06-09 10:52:32] AvatarToggle: Active [2025-06-09 10:52:32] Initial AvatarToggle: Active [2025-06-09 10:52:33] AvatarToggle: Hide [2025-06-09 11:06:59] Question: [BLANK_AUDIO] [BLANK_AUDIO] And press A to record press X to send. And press E to cross X to send. [BLANK_AUDIO] [2025-06-09 11:06:59] Response: It seems like you've uploaded some audio files, but I can't process audio directly. If you have any questions about derivatives or need help with math concepts, feel free to ask! [2025-06-09 11:06:59] ResponseTime: 6.51s	[2025-06-09 11:13:20] AvatarToggle: Active [2025-06-09 11:14:00] Question: [BLANK_AUDIO] Can you say something? [BLANK_AUDIO] Hello. Okay. [BLANK_AUDIO] [2025-06-09 11:14:00] Response: Hi there! How can I assist you today? If you have any math questions, especially about derivatives, just let me know! [2025-06-09 11:14:00] ResponseTime: 4.87s [2025-06-09 11:19:49] Question: I think that answer is 0.5. How do you think? [2025-06-09 11:19:49] Response: It sounds like you're discussing a specific problem. Can you share the details of the problem? Then I can help you understand how to find the derivative and check if 0.5 is correct! [2025-06-09 11:19:49] ResponseTime: 7.75s

P18	[2025-06-09 11:52:31] AvatarToggle: Active [2025-06-09 11:52:31] Initial AvatarToggle: Active [2025-06-09 11:52:40] AvatarToggle: Hide [2025-06-09 12:04:58] Question: [BLANK_AUDIO] Hello. [BLANK_AUDIO] [BLANK_AUDIO] [2025-06-09 12:04:58] Response: It seems like there was an audio file uploaded, but I can't process audio. If you have any math questions, especially about derivatives, feel free to ask! [2025-06-09 12:04:58] ResponseTime: 6.42s	[2025-06-09 12:16:42] AvatarToggle: Active [2025-06-09 12:17:13] Question: Hello. [BLANK_AUDIO] [BLANK_AUDIO] [2025-06-09 12:17:13] Response: Hi there! If you have any questions about derivatives or anything else in math, just let me know! [2025-06-09 12:17:13] ResponseTime: 4.70s [2025-06-09 12:22:34] Question: How can I function h prime minus minus 2? From at the points minus 2 minus 4, to, through the points 1.5. [2025-06-09 12:22:34] Response: It sounds like you're asking about finding the derivative of a function, possibly at specific points. To find h prime at those points, you need to differentiate the function h. If you provide the function h, I can help you find its derivative and evaluate it at the points you mentioned! [2025-06-09 12:22:34] ResponseTime: 5.43s [2025-06-09 12:23:14] Question: How can I function h prime minus minus 2? From at the points minus 2 minus 4, to, through the points 1.5. [SOUND] [SOUND] [2025-06-09 12:23:14] Response: It seems like there might be some confusion in your question. If you're looking to find the derivative of a function h at specific points, please share the function h. Then, I can help you find h prime and evaluate it at the points you're interested in! [2025-06-09 12:23:14] ResponseTime: 5.35s

7.2 Appendix B: Questionnaire Survey

Questionnaire survey of the research 研究のアンケート調査

We collect your email address solely for the purpose of sending an Amazon gift (or allowance) after the study is completed. Your information will be kept confidential and used only for this purpose.
本調査終了後に Amazon ギフト（または謝礼）をお送りするために、メールアドレスを収集しております。ご提供いただいた情報は本目的のみに使用され、厳重に管理されます。

* Indicates required question

1. Participant number *

2. Age 年齢 *

3. Gender 性別 *

Mark only one oval.

- ◯ Female 女性
- ◯ Male 男性
- ◯ Non-binary ノンバイナリー・その他
- ◯ Prefer not to say 回答したくない

4. VR Experience VR経験 *

Mark only one oval.

- ◯ Never / 使ったことがない
- ◯ Once or twice / 一度か二度
- ◯ Sometimes for entertainment / 娯楽目的で数回
- ◯ Regularly / 定期的に使用
- ◯ Research or development / 研究や開発で使用

5. Math level and confidence 数学レベルと自信 *

Mark only one oval.

- ◯ Only basic arithmetic / 四則演算レベル
- ◯ Algebra & geometry only / 代数・幾何まで
- ◯ Somewhat confident with calculus / 微積分にある程度自信あり
- ◯ Confident in advanced math / 高度な数学にも自信あり
- ◯ Math major or frequent use / 数学専攻・常に使っている

6. English Proficiency Level 英語能力レベル *

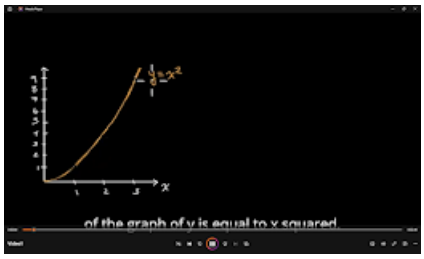
Mark only one oval.

- ◯ 基礎レベル（TOEIC < 550 など）
- ◯ 中級レベル（TOEIC 550-785 など）
- ◯ 上級レベル（TOEIC 785-900 など）
- ◯ 熟練レベル（TOEIC 900+ など）

Condition Select 条件選択

7. Please select the condition you experienced. *
体験した条件を選んでください

Mark only one oval.



Condition 1: PC Viewing / パソコンで視聴 *Skip to question 8*



Condition 2: VR without AI Avatar / VRでAIアバターなし *Skip to question 15*



Condition 3: VR with AI Avatar / VRでAIアバターあり *Skip to question 22*



All completed 全部完成 *Skip to question 33*

Condition 1: PC Viewing / パソコンで視聴

Condition 1: Only watch the instructional video in the computer.
条件1:インストラクションビデオをパソコンで見ること。

8. Immersion *
没入感

I felt immersed in the experience.
私はこの体験に没入感を感じました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

9. Enjoyment *
楽しさ

I enjoyed the experience.
私はこの体験を楽しみました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

10. Comprehension *
理解度

I understood the content of the experience well.
私は体験の内容をよく理解できました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

11. Mental Demand *
精神的負荷

The experience was mentally demanding.
この体験は精神的な負荷が高かったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

12. Physical Demand *
身体的負荷

The experience was physically demanding.
この体験は身体的な負荷が高かったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

13. Effort *
努力度

I had to put a lot of effort into this experience.
この体験には多くの努力が必要でした。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

14. Frustration *
フラストレーション

I felt frustrated during the experience.
私はこの体験中にフラストレーションを感じました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

Skip to question 7

Condition 2: VR without AI Avatar / VRでAIアバターなし

15. Immersion *
没入感

I felt immersed in the experience.
私はこの体験に没入感を感じました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

16. Enjoyment *
楽しさ

I enjoyed the experience.
私はこの体験を楽しみました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

17. Comprehension *
理解度

I understood the content of the experience well.
私は体験の内容をよく理解できました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

18. Mental Demand *
精神的負荷

The experience was mentally demanding.
この体験は精神的な負荷が高かったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

19. Physical Demand *
身体的負荷

The experience was physically demanding.
この体験は身体的な負荷が高かったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

20. Effort *
努力度

I had to put a lot of effort into this experience.
この体験には多くの努力が必要でした。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

21. Frustration *
フラストレーション

I felt frustrated during the experience.
私はこの体験中にフラストレーションを感じました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

Skip to question 7

Condition 3: VR with AI Avatar / VRでAIアバターあり

22. Immersion *
没入感

I felt immersed in the experience.
私はこの体験に没入感を感じました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

23. Enjoyment *
楽しさ

I enjoyed the experience.
私はこの体験を楽しみました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

24. Comprehension *
理解度

I understood the content of the experience well.
私は体験の内容をよく理解できました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

25. Mental Demand *
精神的負荷

The experience was mentally demanding.
この体験は精神的な負荷が高かったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

26. Physical Demand *
身体的負荷

The experience was physically demanding.
この体験は身体的な負荷が高かったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

27. Effort *
努力度

I had to put a lot of effort into this experience.
この体験には多くの努力が必要でした。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

28. Frustration *
フラストレーション

I felt frustrated during the experience.
私はこの体験中にフラストレーションを感じました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

Avatar Evaluation (Condition 3 Only) / アバター評価（条件3のみ）

Please select the option in Condition 3 (VR With Avatar) that best matches your feelings of the Avatar.
条件3(アバター付きVR)でアバターの気持ちに一番合ったものを選んでください。

29. Human-likeness *
人間らしさ

The AI avatar was human-like.
AIアバターは人間のようでした。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

30. Usefulness *
実用性

The AI avatar was useful in the experience.
AIアバターはこの体験において役に立ちました。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

31. Pleasantness *
快適さ

The AI avatar was pleasant.
AIアバターは好ましかったです。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

32. Appearance *
外見

I liked the AI avatar's appearance.
私はAIアバターの見た目が好きでした。

Mark only one oval.

	1	2	3	4	5	6	7	
Stro	◯	◯	◯	◯	◯	◯	◯	Strongly Agree とてもそう思う

Skip to question 7

User Preference

33. Please rank the following learning experiences from your most preferred (**A**) to least preferred (**C**): *
以下の学習経験を、あなたが最も好む(**A**)から最も好まない(**C**)までランク付けしてください。

Mark only one oval per row.

	A (most preferred 最も好む)	B (general 普通)	C (least preferred 最も好まない)
Condition 1: PC Viewing / パソコンで視聴	◯	◯	◯
Condition 2: VR Only / VRだけ	◯	◯	◯
Condition 3: VR with AI Avatar / VRでAIアバターあり	◯	◯	◯

34. Why did you choose to rank this condition as A (most preferred)? *
この条件をA（最も好む）として選んだのはなぜですか？

35. Why did you choose to rank this condition as C (least preferred)? *
この条件をC（最も好まない）として選んだのはなぜですか？

This content is neither created nor endorsed by Google.

Google Forms

7.3 Appendix C: Quiz for Research

Quiz for Research

* Indicates required question

1. Participant number *

2. Quiz selection クイズに選択 * Dropdown

Mark only one oval.

- Quiz 1 (Video 1) *Skip to question 3*
- Quiz 2 (Video 2) *Skip to question 4*
- Quiz 3 (Video 3) *Skip to question 5*

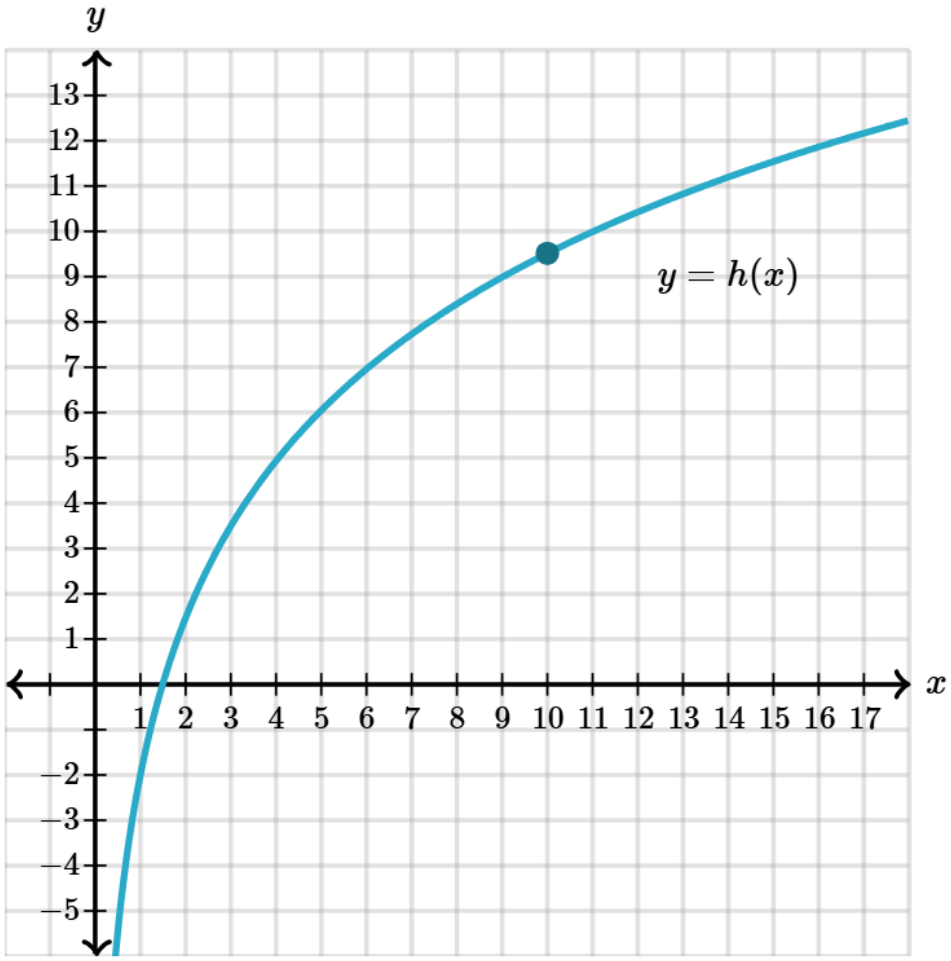
Quiz 1

3. What is the average rate of change of g(x) = 2x over the interval [3, 10]? *

Mark only one oval.

- 2
- 1
- 3.5
- 14
- 7

Quiz 2



4. Estimate h'(10). *

Mark only one oval.

- 0.5
- -10
- 0
- -0.5
- 10

Quiz 3

5. The tangent line to the graph of function h at the point (-2,-4) passes through the point (1,5). *
Find h'(-2).

Mark only one oval.

- ◯ -3
- ◯ 2
- ◯ 3
- ◯ 1.5
- ◯ 9

This content is neither created nor endorsed by Google.

Google Forms

7.4 Appendix D: Feedback

	Interview
P1	- The existence of Avatar will make me feel more secure when solving quizzes. - But the quizzes are too simple, so I don't need to interact with it.
P2	- The video quality in VR doesn't look very good.(blur)
P3	- It feels like avatar is highly practical in any scene. - I hope to increase multilingual interaction. - The Avatar and the video content are not synchronized. I don't want to repeat the quiz questions. I hope it can be as convenient and fast as Chatgpt.
P4	- The distance between the Avatar and the blackboard is too far.
P5	- It's too fast for non-native English speakers. I want to adjust the speaking speed of the video. - I hope there will be more interactions, such as adding translation functions and automatically generating multilingual subtitles.
P6	- Watching videos in VR is just like being in a real classroom - VR is highly immersive and makes people more focused. - I think the experience of Condition2 and Condition3 is the same, so it need to further leverage the characteristics and features of VR.
P7	- VR is very strong sense of immersion and is very close to reality. Directly communicating with the Avatar is better than using a computer. - The mathematic quizzes felt a bit difficult. Because it was in English, it seemed even more challenging (for non-native speakers). - Avatar is helpful. Its presence is attractive, so it makes me listen more carefully
P8	- Drafts cannot be used in VR, and VR can only be used to watch videos. - The VR environment has no distractions, which makes me more focused. - This is my first time using VR. I feel very interest - I hope there are draft and notebook functions in the VR environment.
P9	- Avatar looks very much like a teacher, which makes me pay attention - I hope Avatar can also watch the video together and interact with me - I hope there are also draft tools in VR
P10	- I was very surprised by Avatar's response because even if I couldn't express some mathematical functions, it could still guess what you were going to say. - The Video is too long. I don't think it needs that long.
P11	- The VR sense of immersion is very strong, and one can feel as if their entire being is inside - The speaking speed of the Video is too fast. Sometimes it's hard to hear clearly. - Video and Avatar are not interconnected. It is hoped that Avatar can be heuristic. - I think the existence of Avatars would be very useful, but adjustments need to be made according to different scenarios. - I think the more the Avatar looks like a real person, the better. This one is not look that real. - I want an Avatar that can be personalized.
P12	- The interactive interface could be simpler. Some buttons are too small. - Avatar needs to be more proactive in giving hints rather than making users think about how to interact.
P13	- The quizzes are too simple - Everything else is well done
P14	- Condition 1 (computer) makes it difficult to concentrate. - The speaking speed of the video is too fast and not friendly to non-native English speakers. - Avatar looks very intelligent and powerful. - VR environment can significantly enhance the sense of immersion, allowing focus more.
P15	- The classroom environment is great and there is no unnecessary interaction. - Avatar just standing there looks very awkward and has a delayed response.
P16	- Wearing glasses makes me uncomfortable when wearing VR headsets - VR cannot use drafts. - The cable really affects my experience. I hope it could be a little longer - The questions of Quizzes cannot be zoomed in - The speech recognition is inaccurate. I hope to add a keyboard for assistance
P17	- I enjoy it very much, but I don't know the meaning.
P18	- Avatar doesn't look like a human. I hope there is a version that looks more like a human. - The feedback from Avatar is not good. I hope the voice recognition can be more accurate. - Hope to see the current time in the classroom, the clock is fake.