Question-Asking by AI-Driven Virtual Peers Improves Learning Outcomes in VR Classrooms

Category: Research

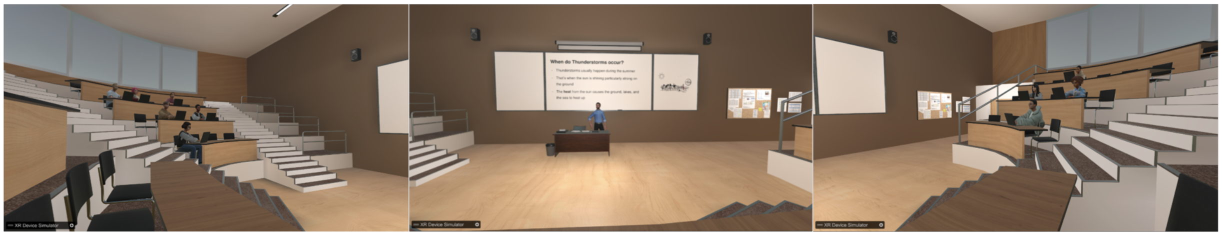


Figure 1: View of the AI classroom as seen from a study participant's VR headset with the central image showing the instructor presenting lecture slides, and the left and right perspectives show the surrounding virtual classroom with AI peers seated to either side.

Abstract

Asynchronous online learning platforms such as Coursera and Khan Academy reach millions of learners but often lack the peer interactions that support engagement and comprehension in traditional classrooms. In this work, we developed a VR classroom where GPT-4-powered virtual peers pose contextual, domain-relevant questions during slide-based lectures. These AI-generated questions are strategically timed to prompt elaboration from the virtual instructor, simulating the benefits of peer inquiry. In a controlled study with 36 participants, we compared two conditions: VR lectures with and without AI questioning agents. Participants in the AI question-asking condition showed significantly higher engagement and focus with improved learning outcomes. These results suggest that simulated peer questioning in immersive VR environments can replicate key benefits of classroom social dynamics, offering a scalable path to enhance online education.

Index Terms: Virtual Reality, AI Peers, Learning, Question-asking

1 Introduction

Active learning engages students in questioning, discussion, and problem solving, and consistently improves comprehension, motivation, and retention compared to passive instruction [27, 11]. One driver of these benefits is peer interaction arising from classmates asking questions, which introduces perspectives a learner may not have considered, adds contextual detail, and models curiosity that encourages further inquiry [8, 9].

Most large-scale online education platforms such as Coursera, Khan Academy, and Udemy rely on asynchronous video lectures, often supplemented with discussion or Q&A forums. These forums provide a form of online social interaction, and frequent forum contributors can become recognized by their peers over time [13]. However, participation is typically optional and reactive such that learners turn to forums often when they encounter difficulties with assignments or concepts, rather than as part of the instruction itself [20]. As a result, peer exchange happens after the fact, disconnected from the flow of instruction. This contrasts with classroom settings, where hearing clarifying questions asked by peers during lectures provides immediate alternative perspectives, additional detail, and models curiosity in the moment of learning [8, 9].

To address this disconnect between instruction and peer interaction, recent technologies offer new possibilities. Virtual reality (VR) can recreate classroom immersion and co-presence, supporting attention and social interaction [28]. Large language models (LLMs) such as ChatGPT [2] can generate contextually appropriate dialogue and role-play as peers or instructors. Prior work has explored VR to increase presence in learning environments [18] and LLMs as personalized tutors or problem-solving assistants [12], but their integration to replicate real-time peer questioning during instruction remains underexplored. Unlike asynchronous forums—where questioning is delayed, optional, and reactive, VR classrooms with LLM-powered peers can embed clarifying questions directly into the lecture. This makes peer interaction continuous, proactive, and accessible to all learners, not just those who post in forums.

In this paper, we investigate whether AI-driven peer questioning in a VR classroom improves learning outcomes relative to a passive VR lecture where no questions are asked. We developed a VR environment in which GPT-powered virtual peers posed clarifying questions during an instructor-led lecture. In a study with 36 participants, we compared this peer-questioning condition to a passive lecture condition, measuring knowledge retention, focus, and engagement. Our findings show that simulated peer questioning can enhance learning in VR, pointing toward scalable methods to embed active learning dynamics into online education.

This work contributes to understanding how AI-driven peer dynamics can support active learning in VR. Specifically, our contributions include:

- Empirical evidence that AI-generated peer questioning improves learning outcomes in a VR classroom compared to passive VR lectures, with measurable gains in learning outcomes, focus, and engagement.
- Evidence that real-time peer questioning by simulated peers can reproduce classroom interaction benefits, with potential to scale this approach to different domains, instructional formats, and expanded VR classroom designs.
- Design implications for scalable online education, highlighting how VR and LLMs together can embed active learning mechanisms into digital environments at scale.

2 Related Work

Research on online and virtual learning has explored multiple ways to enhance engagement and social presence, ranging from interactive instructors to peer evaluation, embodied pedagogical agents, and AI-generated classmates. Here, we summarize areas most relevant to our work.

2.1 Active vs Passive Instruction

Active learning refers to instructional approaches that involve students directly in the learning process through meaningful activities and reflection that promote participation and cognitive effort [27]. In contrast, traditional online learning environments often rely on passive instruction, where students consume prerecorded content without real-time interaction or feedback. Research shows that students learn more effectively when they actively participate rather than passively receive information [27, 11]. In classroom contexts, this participation often takes a social form, such as peer discussions, collaborative problem solving, or asking and answering questions during lectures. These dynamics not only require cognitive engagement but also expose learners to alternative perspectives and model curiosity in real time [8, 10].

Although online learning increases accessibility, passive formats pose persistent challenges, including limited instructor interaction, cognitive overload, reduced motivation and focus, language-related difficulties, and struggles with managing time effectively [24]. Active learning strategies, by contrast, consistently improve conceptual understanding, with particularly strong effects in small classes [27, 11]. These findings indicate the importance of embedding interactive and participatory elements into digital learning environments, rather than relying solely on content delivery.

One technological approach to implementing these principles has been the use of conversational agents. Such agents enable two-way interaction, provide immediate feedback, and foster more socially engaging forms of online education [26, 16, 7]. Studies comparing passive virtual instructors to interactive AI-driven instructors in web-based platforms consistently report gains in motivation, engagement, and perceived instructor humanness [25, 36]. However, these improvements are often limited to subjective experience and do not always translate into measurable performance gains.

Beyond text- and video-based platforms, virtual reality (VR) has also been investigated as a medium that can recreate immersion and co-presence, supporting attention and social interaction [28, 34]. VR classrooms with interactive agents extend opportunities for embodied engagement, where co-presence may amplify the perceived authenticity of social interaction. Yet, prior work in this space has primarily focused on interactive instructors, leaving the role of simulated peers underexplored.

Our study builds on these threads in two ways. First, we shift attention from the instructor to the peer role, examining whether simulated peers can enhance learning through questioning. Second, while earlier research has largely examined instructor interactivity in 2D web environments, we embed our experiment in immersive VR, where co-presence may further amplify the effects of peer interaction.

2.2 Peer Evaluation

Although instructors are often assumed to be the most qualified evaluators in education, research shows that students themselves can play a valuable role in the learning process. Peer evaluation has been found to produce consistent grading and to provide both cognitive and motivational benefits [9, 35]. Several studies demonstrate that student peer assessments, when guided by clear rubrics and accountability, can approach instructor assessments in reliability and validity [30, 32], challenging traditional assumptions about instructor authority. Later studies emphasize that peer evaluation exposes learners to diverse perspectives, fosters reflection, and deepens understanding through engagement with work by others [33, 23].

Together, this body of research highlights the broader value of incorporating peer contributions into learning for accountability, modeling curiosity, and introducing alternative perspectives. Building on this foundation, our study shifts the focus from evaluation to questioning. Instead of grading, our virtual peers ask clarifying questions during a lecture. By situating peer contributions within the flow of instruction rather than after the fact, we test whether simulated peer questioning can yield similar benefits of perspective-taking and reflection in VR-based classrooms.

2.3 Animated Pedagogical Agents

Animated pedagogical agents (PAs) have been widely studied as a means of increasing social presence in multimedia learning environments. Through gestures, gaze, and speech, PAs simulate aspects of human communication that encourage deeper engagement from learners [6]. For example, Li et al's [17] study on chemical synaptic transmission found that learners who studied with an embodied PA not only performed better on retention and transfer tests, but also showed increased activation in brain regions associated with social cognition and observational learning, as measured by fNIRS. These findings align with social agency theory [22], which holds that learners process information more deeply when they perceive an agent as a social partner.

Prior research on PAs has focused primarily on the instructor role, where agents act as teachers or guides. In contrast, our study examines whether similar social-cognitive benefits can arise from peer agents. We simulate virtual students who ask clarifying questions during VR lectures, extending research on social presence beyond instructors to peers. This approach allows us to test whether human-like behaviors from student agents can activate the same social processing mechanisms and enhance learning outcomes.

2.4 Interactive Virtual Peers in VR

The most directly related work to ours is ClassMeta [19], which used GPT-powered agents to simulate classmates in VR. These virtual peers engaged in a range of behaviors, including note-taking, asking questions, responding to instructor prompts, promoting discussions, and even correcting off-task behavior. Results showed that learners maintained sustained attention to the agents and produced higher-quality notes and stronger reasoning outcomes, suggesting that AI-generated classmates can model constructive behaviors and support social learning.

While ClassMeta demonstrated the promise of VR classrooms populated with interactive peers, its agents performed many behaviors simultaneously, making it difficult to isolate which mechanisms were most impactful. Our study narrows this focus to a single key behavior: peer questioning. By examining the specific contribution of clarifying questions, we provide a more fine-grained understanding of how peer agents influence motivation, comprehension, and engagement. In this way, we build directly on ClassMeta while contributing evidence about the distinct pedagogical value of simulated peer questioning.

Together, prior work highlights the importance of interactivity, social presence, and peer contributions in shaping learning experiences. Yet most studies have focused on instructors, delayed forms of peer input such as evaluation or discussion forums, or multi-behavior peer agents whose pedagogical mechanisms are difficult to disentangle. Our work addresses these gaps by isolating and testing the role of peer questioning in a VR classroom, examining its effects on motivation, comprehension, and engagement. In doing so, we extend active learning research into immersive environments and demonstrate the distinct pedagogical value of simulated peer dynamics.

3 System Design

The system is designed to generate real-time audio interactions in VR after each lecture slide, using the OpenAI API. After parsing slide text, a script constructs a prompt using the parsed information and generates an audio file using OpenAI for the virtual instructor. The response given by the instructor is passed back to OpenAI to generate a question that the virtual student can ask. The instructor then have an audio file generated in realtime as a response to

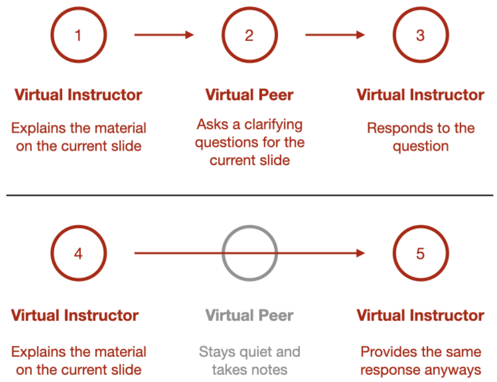


Figure 2: Study design showing Active and Passive conditions. In the Active condition (top), a random AI peer asks a question after each slide, prompting instructor elaboration. In the Passive condition (bottom), the instructor delivers identical elaborative content without AI questioning. Both conditions provide equivalent instructional material while isolating the effect of peer-initiated question-asking.

this question. This pipeline is intended to generalize across lecture materials, allowing any slide deck to be automatically transformed into an interactive VR lecture. In the full design, AI-driven peers can ask clarifying questions and the AI-driven instructor responds in real time. However, during pilot testing the model often produced verbose or filler-laden responses, substantially extending lecture duration. In addition, the peer questions varied across runs, introducing inconsistencies between sessions.

To balance realism with experimental control, we modified the design. Our goals in constructing the VR classroom were threefold: (1) to create a realistic and distraction-free environment that supported a sense of co-presence, (2) to implement AI avatars and behaviors that modeled classroom dynamics such as listening, questioning, and note-taking, and (3) to control timing, voice, and animation so that interactions were consistent across participants. While our architecture supports both AI-driven peers and an AI-driven instructor who responds to their questions, for the user study we scripted and synchronized all interactions to maintain uniformity across conditions. These choices ensured both ecological validity for learners and the experimental control required for evaluation.

3.1 Platform and Hardware

The VR classroom was developed in Unity [4] using the XR Plugin Environment. The base environment was adapted from the Unity Asset Store University Classroom asset [1]. To minimize distractions, extraneous visual elements were removed, and laptops were positioned at each student seat to signal participation. The projector screen was scaled for clarity and brightness, ensuring lecture content was easily legible. All dimensions were adjusted to a 1:1 scale with avatars and participants to enhance spatial realism.

3.2 Virtual Teacher and Student Avatars

Virtual students were instantiated as ReadyPlayerMe avatars [5] and distributed throughout the classroom. To simulate classroom behavior, we incorporated Mixamo[1] animations for sitting, typing, hand-raising, talking, and pointing. Animations were triggered through state machines with Boolean conditions, allowing avatars to shift naturally between attentive listening, asking a question, and responding. These behaviors provided nonverbal cues to reinforce the perception of social co-presence.

3.3 Audio Generation and Synchronization

Lecture delivery was generated from slide content using a three-part pipeline: (1) teacher explanation of each slide, (2) a clarifying student question, and (3) a teacher response (Figure 3). Prompts constrained the outputs to remain short, accurate, and accessible ($\leq$ 60 words for explanations, $\leq$ 50 words for answers, and single-sentence clarifying questions). This ensured a consistent rhythm of instruction and peer-style questioning without introducing off-topic material.

For each textual response, male and female audio files were synthesized using OpenAI's TTS API [2]. A C# script synchronized playback with avatar positioning and animations: explanations were paired with teacher gestures, while student questions were accompanied by hand-raising and talking animations. Teacher gender was systematically varied by participant group for experimental control, while student questioners were randomly assigned male or female voices to create a balanced mix.

4 Study Design

To evaluate the impact of AI-driven peer questioning in virtual classrooms, we conducted a controlled between-subjects experiment with two conditions: a passive VR lecture and an active VR lecture featuring virtual students who asked clarifying questions (Figure 2). All sessions were delivered on Meta Quest 3 headsets [3] in a quiet laboratory setting.

4.1 Participants

Thirty-six participants (33 undergraduate students and 3 recent graduates) were recruited from a university campus. None reported prior expertise with the lecture topic, *Thunderstorms Explained*. Ages ranged from 18–25 (10F, 26M). Participants received \$10–\$15 compensation (depending on the length of the study, which was 2 minutes longer for those in the Active condition due to AI peer questions) via Zelle or Venmo.

4.2 Experimental Design and Conditions

The study used a between-subjects design with two conditions:

Active Condition (Experimental): Participants attended a VR lecture, presented slide by slide. Virtual (AI) students were placed throughout the classroom. After each slide, the instructor asked the class if there were any questions. A random virtual student raised their hand, prompting the instructor to turn, point, and call on them. The student then asked a clarifying question related to the slide, repeating words from the slide without adding new information, which the instructor answered before proceeding to the next slide (Figure 3). The peer questions were intentionally hard-coded to avoid introducing any new information, ensuring that participants in the active lecture did not receive an unfair informational advantage. Question-asking added two minutes to this condition over the Passive condition.

Passive Condition (Control): Participants attended the same VR lecture as in the Active condition, delivered by a virtual instructor with no questions asked by virtual students. After each slide, the instructor asked the class if there were any questions. However, in the passive lecture, no students raised their hands. The instructor still continued as if someone had asked a question, reiterating the material on the slide with the same response used in the Active condition. The instructor then proceeded to the next slide and repeated this process throughout the lecture. This version was structured to simulate a classroom with student attendance but no student participation.

To ensure valid comparisons across conditions, several controls were implemented. The lecture topic was selected for its moderate

[1] https://www.mixamo.com/

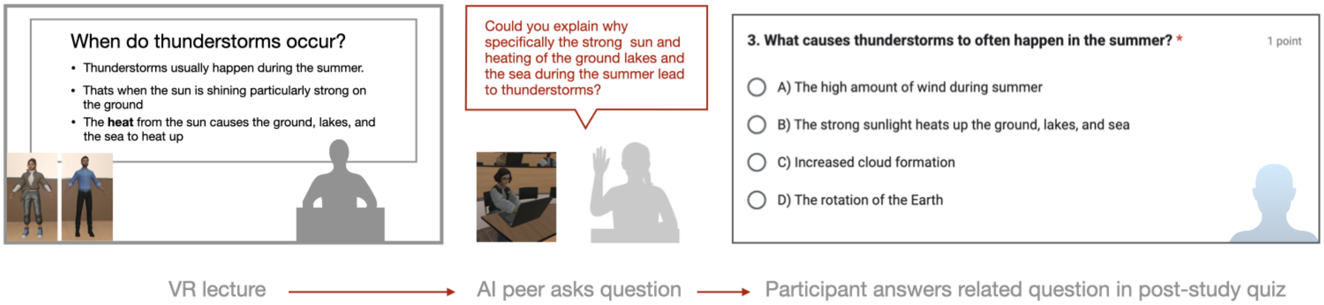


Figure 3: Illustration of the participant experience in the Active condition: (left) instruction content delivered in the VR lecture, (middle) an AI peer asking a clarifying question related to the slide, and (right) a corresponding quiz item assessing understanding of the same material. Insets show the two VR instructors and the AI peer as seen in VR.

complexity and disciplinary neutrality, based on prior use in educational research by Mayer et al.[14, 21] (Figure 2). To isolate the effect of peer questioning, participants were instructed not to take notes during the lecture[31]. All instructor and student dialogue was pre-generated using GPT-4 and OpenAI's text-to-speech system to maintain consistent content and timing [2]. The lecture was presented in a fixed sequence of audio and animations, minimizing variance while simulating a realistic delivery. The information the instructor shared in response to the AI peer's question in the Active condition was also used as part of the lecture for the Passive condition, ensuring that participants in both conditions received the same information. However, a participant in Active lecture spent a little more time in the classroom, owing to the time it took the virtual peer asking the question. All participants experienced the same VR classroom environment, including uniform lighting, seating, and screen visibility.

Participants were randomly assigned a participant ID from 1 to 36. Group assignment was dependent on the ID number with odd-numbered IDs in the Passive condition and even-numbered IDs in the Active condition. Instructor gender was counterbalanced across participants, virtual peer questioners were randomly picked from the available auditorium seats and assigned male or female voices. Dialogue was pre-generated using GPT-4 and OpenAI TTS [2] to ensure consistent delivery across participants.

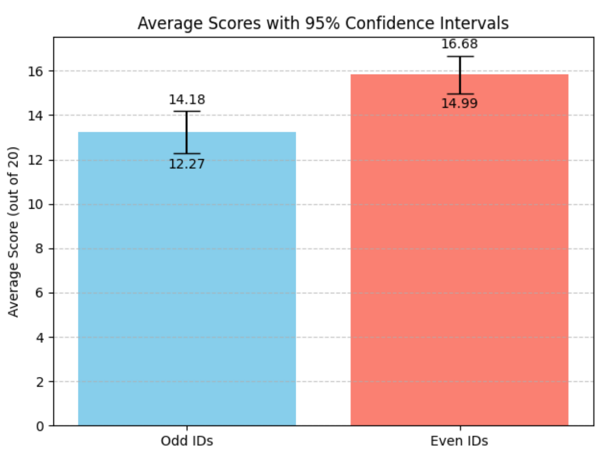


Figure 4: Average post-study group scores with 95 CI are significantly higher in the Active condition (Even IDs) compared to the Passive condition (Odd Ids).

4.3 Procedure

Each study session lasted approximately 40–45 minutes and followed a consistent sequence across participants. After providing informed consent (study approved by our IRB under protocol # anonymous), participants completed a short pre-survey that collected demographic information and assessed prior familiarity with thunderstorms. They then put on a Meta Quest 3 headset [3] and experienced either the passive or active lecture condition, which lasted 7–9 minutes. Immediately afterward, they completed a post-lecture knowledge assessment consisting of multiple-choice and short-answer questions designed to test both retention and understanding of the material. This was followed by a post-survey that measured engagement, focus, perceived learning and effectiveness on 5-point Likert scale questions. To conclude, participants filled out a motion sickness questionnaire to report any discomfort experienced in VR.

4.4 Measures

We assessed learning outcomes through performance on a post-lecture quiz (Figure 3). Engagement and motivation were measured via self-reports in a custom post-study survey. User experience was captured through ratings of satisfaction, focus, and perceived learning effectiveness from the IPQ questionnaire [29]. Physical discomfort and motion sickness were evaluated using the SSQ questionnaire [15].

Learning outcome was measured with a quiz consisting of 20 multiple-choice questions. Each question offered four options with a single correct answer. Items tested both knowledge retention and knowledge transfer, with 10 questions for each of these two groups. For example: *"How does lightning produce thunder?"* was a question in the knowledge retention section, and *"How might increased humidity affect the likelihood of lightning?"* was a question in the knowledge transfer section. Each question had a value of 1 point, so the maximum score for the quiz was 20 points.

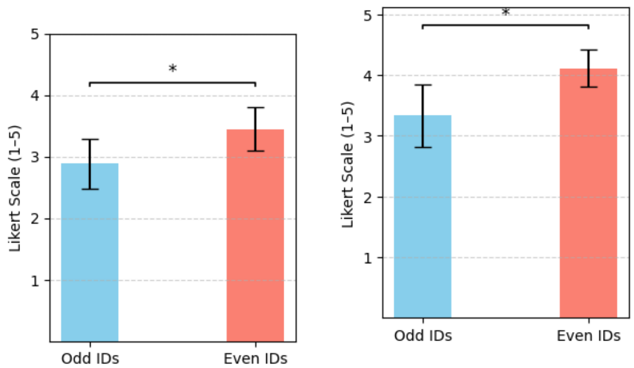
5 Results

5.1 Post-Study Quiz Performance

The post-lecture assessment included retention (recall) and transfer (application) items. See Figure 4.

Assumption Checks. A Shapiro–Wilk test indicated no significant deviation from normality for either group ($p > .05$). Levene's Test confirmed homogeneity of variance, $F(1,34) = 0.413$, $p = .525$.

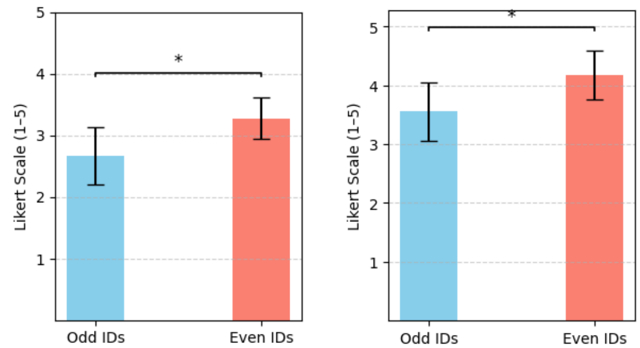
Group Comparison. Participants in the Active condition (M = 15.8, 95% CI [14.99, 16.68], SD = 1.89) scored significantly higher than those in the Passive condition (M = 13.2, 95% CI [12.27, 14.18], SD = 2.13). The mean difference of 2.6 points (out of 20)



(a) I felt actively engaged throughout the lecture (p = 0.036).

(b) The presence of other students made the classroom feel more realistic (p = 0.018).

Figure 5: Significantly higher engagement and sense of social presence. Odd IDs correspond to the Passive group and Even IDs to the Active group of participants.



(a) The classroom's social environment helped me stay focused (p = 0.021).

(b) I found the instructors delivery to be clear and understandable (p = 0.040).

Figure 6: Significantly higher focus and perception of clarity of instruction. Odd IDs correspond to the Passive group and Even IDs to the Active group of participants.

corresponds to a 19.74% improvement for the Active group. A two-tailed t-test confirmed this effect, $t(34) = 2.71$, $p < .001$, with a large effect size (Cohen's $d \approx 0.9$).

Association. A point-biserial correlation indicated a moderate-to-strong relationship between study condition and quiz score, $r = .56$, $p < .001$.

5.2 Engagement and Learning Experience Questionnaire

To complement the objective performance data, participants completed a custom post-study Likert-scale questionnaire that assessed engagement, focus, and perceived learning during the VR lecture. Five core questions revealed statistically significant differences between the Passive and Active conditions, with the Active group consistently reporting higher scores. Mann–Whitney U tests were used for analysis, as appropriate for ordinal data.

Results indicated that the Active group felt more engaged in the lecture ($U = 215.0, p = 0.036$) as seen in Figure 5aa), perceived the presence of other students as making the classroom feel more realistic ($U = 225, p = 0.018$) as shown in Figure 5bb), and reported that the classroom's social environment helped them stay focused ($U = 225.5, p = 0.021$; Figure 6aa). They also more strongly agreed that they learned something new from the session ($U = 209, p = 0.050$; Figure 7aa). Instructor clarity showed a positive and significant trend in favor of the Active group ($U = 215.0, p = 0.040$; Figure 6b). Preference for online video lectures to be delivered in this VR format was significantly higher in the active group than the virtual group ($U = 212.50, p = 0.046$; Figure 7b).

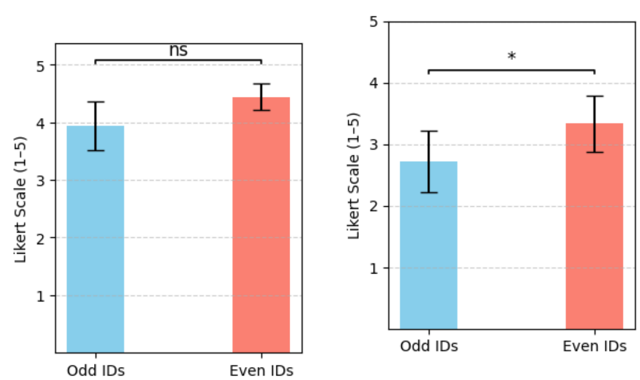
There was no significant difference in pacing of the lecture feeling appropriate ($U = 207.00, p = 0.069$), but both groups averaged above a 3 out of 5 on the Likert scale. Enjoyment of the virtual classroom experience was not significantly different either ($U = 164.50, p = 0.472$)). There was no significant difference in the the virtual classroom format helping participants understand the topic better than a traditional video ($U = 187.50 p = 0.204$). Finally, there was no significant difference between groups in recommending this experience to a classmate ($U = 177.50, p = 0.3055$). In response to the statement "I used more of my intuition rather than what I actually learned from the class" ($U = 11.0 p = .074$), there was no significant difference between the groups.

5.3 Presence Questionnaire Results

To assess immersion and social presence, participants completed relevant questions in the IPQ questionnaire [29]. While most items showed no significant differences between groups, one key effect emerged. The question "How aware were you of the surrounding real world while navigating in the virtual world" ($U = 103.0$, $p = .028$; Figure 8) scored significantly lower in the Active group than in the Passive group. This showed that participants in the Active condition reported a stronger sense of immersion, feeling less aware of the real world surroundings.

The question "How real did the virtual world seem to you" ($U = 187.5$, $p = .813$) did not show a significant increase between groups. Both groups responded with an average score slightly below 3, showing that there is room for improvement in the realism of the simulation. The question "I was completely captivated by the virtual world" ($U = 164.0$, $p = .533$) did not show a significant increase between groups either. This may seem related to the previous question, since a world that seems less real may be less captivating.

The question "I had a sense of acting in the virtual space, rather than operating in a virtual space" ($U = 171.0, p = .63$) had averages above 3 in both groups. Although this result was not significantly increased in the active group, it shows that both groups had some sense of being embodied within the virtual classroom. The question



(a) I feel like I learned something new from the class session (p = 0.050)

(b) I would prefer more online video lectures to be delivered in this format (p = 0.046).

Figure 7: Significantly higher preference for more lectures delivered in this format. Odd IDs correspond to the Passive group and Even IDs to the Active group of participants.

"I felt present in the virtual space" ($U = 168.0$, $p = .595$) also did not increase significantly between groups, but showed a positive trend in both.

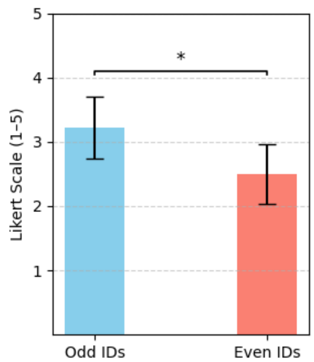


Figure 8: How aware were you of the real world surroundings while navigating the virtual world. Odd IDs correspond to the Passive group and Even IDs to the Active group of participants.

5.4 Active Group Specific Questions

To assess experience of peer-style questioning, we administered four additional items to participants in the Active condition. The participants rated them on a Likert scale from 1 to 5. Responses indicated that participants generally evaluated the virtual questions positively. The statement "Hearing questions from virtual students helped me stay focused during the lecture" received an average rating of 4.06 (SD = 0.91). Participants rated "The questions asked by virtual students helped clarify the material for me" at 4.00 (SD = 1.0), and "The questions asked by virtual students felt relevant to my learning" at 4.11 (SD = 0.87). The last item was "I was thinking about asking similar questions as the other students" with a mean of 3.05 (SD = 1.27).

5.5 Motion Sickness

To evaluate the physical comfort of the VR setup, participants completed the SSQ motion sickness questionnaire at the end of the session. Results revealed no significant differences between the Passive and Active groups using the Mann-Whitney U test, and 35 out of 36 participants verbally reported they were feeling okay at the end of the study.

Table 1: Group comparisons (Odd vs. Even participant IDs) for SSQ composite scores. Values are group means; Mann–Whitney U statistics and p-values test group differences.

Metric	**Odd Mean**	**Even Mean**	**U**	***p*-value**
Nausea	7.16	7.95	154.0	0.780
Oculomotor	25.27	27.16	153.5	0.796
Disorientation	23.20	27.84	140.0	0.469
Total Score	21.51	24.00	147.0	0.642

6 Discussion

Participants in the Active condition demonstrated clear learning benefits over the Passive condition. Quiz performance was nearly 20% higher, with a large effect size, showing that exposure to peer questions improved both recall and transfer. This suggests that questions likely prompted deeper processing of lecture material, rather than simply extending the time spent in VR. Consistent with this interpretation, the Active group also reported higher engagement, stronger focus supported by the classroom's social environment, and a greater sense of realism due to the presence of peers. They further agreed more strongly that they learned something new, rated the instructor as clearer, and expressed a stronger preference for receiving video lectures in this VR format. Together, these results indicate that simulated peer questioning provided meaningful pedagogical value in addition to the baseline VR lecture experience.

At the same time, several measures showed no significant differences between conditions. Both groups rated pacing as appropriate, enjoyment of the VR classroom positively, and comprehension relative to traditional video at similar levels. They were also equally likely to recommend the experience to a classmate and to report relying on intuition rather than learned material. These null results are consistent with the study design: aside from peer questioning, both conditions shared the same VR setting, instructor delivery, pacing, and content. Thus, differences should be expected primarily in outcomes directly related to peer contributions such as focus, clarity, and knowledge retention rather than in overall impressions of the VR platform.

The presence results suggest that peer questioning selectively influenced immersion without broadly shifting other aspects of presence. Participants in the Active group reported being less aware of the surrounding real world, indicating deeper immersion during the lecture. However, no significant differences were found for measures of realism, captivation, or overall presence, with both groups rating these dimensions around or slightly above the midpoint of the scale. This pattern is consistent with the study design: since both conditions used the same VR classroom environment, differences would be expected primarily in attentional focus rather than in perceptions of graphical realism or embodiment.

Responses from the Active group confirm that participants generally valued the peer-style questions. Ratings above 4 for focus, clarification, and relevance indicate that the questions were perceived as supportive of learning. The lower average on the item about thinking of the same questions (M = 3.05) suggests that the benefit did not rely on exact alignment with a participant's own thoughts; rather, the questions were helpful even when they did not directly match what participants themselves would have asked.

Motion sickness was minimal across both conditions. SSQ scores showed no significant differences between Active and Passive groups, and nearly all participants reported feeling comfortable at the end of the study. This indicates that the addition of peer questioning did not introduce extra physical strain and that the VR setup was generally well tolerated.

Overall, the findings show that peer questioning enhanced learning and engagement while leaving the overall acceptability of the VR classroom unchanged. Even with scripted interactions, the presence of peers influenced attention and comprehension in measurable ways across both performance and self-report. These results point to AI-mediated peer dynamics as a focused mechanism for bringing active learning benefits into immersive environments.

7 Limitations & Future Work

This study has several limitations. The sample consisted primarily of university students from a single institution. While this aligns with the classroom learning context, it narrows generalizability to other learner populations such as younger students, adult learners outside academia, or participants with less exposure to technology. The intervention was also limited to a single lecture topic and a short session, leaving open whether the effects extend across domains or longer time scales.

Methodologically, peer interactions were pre-generated to ensure consistency. While this provided strong experimental control, it reduced spontaneity and limited adaptability to individual learner needs. The peer agents only asked clarifying questions, whereas real students engage in a broader range of behaviors such as elaboration, critique, and discussion [9].

Although motion sickness was minimal, VR fatigue can be a potential barrier for extended learning contexts. More natural voice generation, improved range of peer animations, and varied classroom environments may help increase realism, and reduce fatigue.

Future work should expand in several directions. At the study design level, larger and more diverse samples are needed, including participants across majors, balanced gender representation, and individuals with prior VR experience to mitigate novelty effects. Incorporating participant-generated questions would also allow direct comparison with AI-generated ones. On the system side, adaptive personalization such as tailoring the timing and type of questions to a learner's evolving knowledge, represents a key opportunity. Additional studies should also examine instructor–peer dynamics and evaluate long-term outcomes such as knowledge retention and motivation.

8 Conclusion

The study demonstrates that AI-driven peer questioning in VR classrooms can meaningfully enhance both learning outcomes and the subjective classroom experience. Participants who attended lectures with virtual peers asking clarifying questions achieved higher quiz scores, reported greater engagement, and experienced stronger feelings of presence than those in a passive lecture condition. These results suggest that even simple, non-personalized peer interactions can replicate important aspects of active learning, pointing toward new opportunities for scalable, socially enriched online education. As VR and large language models continue to mature, integrating AI-mediated peers offers a promising path for designing immersive learning environments that balance accessibility with the cognitive and social benefits of classroom interaction.

References

[1] University classroom - unity asset store. `https://assetstore.unity.com/packages/3d/props/interior/university-classroom-86107`. Accessed: 2025-06-14. 3

[2] Introducing chatgpt. `https://openai.com/blog/chatgpt`, 2023. Accessed: 2025-06-14. 1, 3, 4

[3] Meta quest 3. `https://www.meta.com/quest/quest-3/`, 2023. Accessed: 2025-06-14. 3, 4

[4] Unity engine. `https://unity.com/`, 2023. Accessed: 2025-06-14. 3

[5] Ready player me. `https://readyplayer.me/`, 2025. Accessed: 2025-06-14. 3

[6] S. Baceviciute, A. Mottelson, T. Terkildsen, and G. Makransky. Investigating representation of text and audio in educational vr using learning outcomes and eeg. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020. 2

[7] R. A. Carter Jr, L. Zhang, T. L. Hunt, C. R. Emerling, S. Yang, and J. Rujimora. Conversational agents to support remote personalized instruction for diverse learners. *TechTrends*, 67(4):626–636, 2023. 2

[8] M. T. Chi. Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in cognitive science*, 1(1):73–105, 2009. 1, 2

[9] R. J. Dufresne, W. J. Gerace, W. J. Leonard, J. P. Mestre, and L. Wenk. Classtalk: A classroom communication system for active learning. *Journal of Computing in Higher Education*, 7:3–47, 1996. 1, 2, 7

[10] R. J. Dufresne, W. J. Gerace, W. J. Leonard, J. P. Mestre, and L. Wenk. Classtalk: A classroom communication system for active learning. *Journal of computing in higher education*, 7(2):3–47, 1996. 2

[11] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23):8410–8415, 2014. 1, 2

[12] S. García-Méndez, F. de Arriba-Pérez, and M. d. C. Somoza-López. A review on the use of large language models as virtual tutors. *Science & Education*, 34(2):877–892, 2025. 1

[13] J. Huang, A. Dasgupta, A. Ghosh, J. Manning, and M. Sanders. Superposter behavior in mooc forums. In *Proceedings of the first ACM conference on Learning@ scale conference*, pp. 117–126, 2014. 1

[14] C. I. Johnson and R. E. Mayer. A testing effect with multimedia learning. *Journal of Educational Psychology*, 101(3):621, 2009. 4

[15] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993. 4

[16] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori. Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1):973–1018, 2023. 2

[17] W. Li, F. Wang, R. E. Mayer, and T. Liu. Animated pedagogical agents enhance learning outcomes and brain activity during learning. *Journal of Computer Assisted Learning*, 38(3):621–637, 2022. 2

[18] R. Liu, L. Wang, T. A. Koszalka, and K. Wan. Effects of immersive virtual reality classrooms on students' academic achievement, motivation and cognitive load in science lessons. *Journal of Computer Assisted Learning*, 38(5):1422–1433, 2022. 1

[19] Z. Liu, E. Jiang, Z. Zhu, X. Hu, K. Ramani, L. Zhu, and K. Peppler. Classmeta: Designing interactive virtual classmate to promote vr classroom participation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pp. 1–17. ACM, 2024. 2

[20] Z. Liu, N. Zhang, X. Peng, S. Liu, Z. Yang, J. Peng, Z. Su, and J. Chen. Exploring the relationship between social interaction, cognitive processing and learning achievements in a mooc discussion forum. *Journal of Educational Computing Research*, 60(1):132–169, 2022. 1

[21] R. E. Mayer. Cognitive theory of multimedia learning. In R. E. Mayer, ed., *The Cambridge Handbook of Multimedia Learning*, pp. 31–48. Cambridge University Press, 2005. 4

[22] R. E. Mayer, K. Sobko, and P. D. Mautone. Social cues in multimedia learning: Role of speaker's voice. *Journal of educational Psychology*, 95(2):419, 2003. 2

[23] C. Meyers and T. B. Jones. *Promoting Active Learning: Strategies for the College Classroom*. ERIC, 1993. 2

[24] T. T. A. Ngo, T. T. Tran, G. K. An, and P. T. Nguyen. Students' perception towards learning massive open online courses on coursera platform: Benefits and barriers. *International Journal of Emerging Technologies in Learning*, 18(14), 2023. 2

[25] P. Pataranutaporn, J. Leong, V. Danry, A. P. Lawson, P. Maes, and M. Sra. Ai-generated virtual instructors based on liked or admired people can improve motivation and foster positive emotions for learning. In *2022 IEEE Frontiers in Education Conference (FIE)*, pp. 1–9. IEEE, 2022. 2

[26] D. S. Pereira, F. Falcão, L. Costa, B. S. Lunn, J. M. Pêgo, and P. Costa. Here's to the future: Conversational agents in higher education-a scoping review. *International Journal of Educational Research*, 122:102233, 2023. 2

[27] M. Prince. Does active learning work? a review of the research. *Journal of Engineering Education*, 93(3):223–231, 2004. 1, 2

[28] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & education*, 147:103778, 2020. 1, 2

[29] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. 4, 5

[30] C. Schunn, A. Godley, and S. DeMartino. The reliability and validity of peer review of writing in high school ap english classes. *Journal of Adolescent & Adult Literacy*, 60(1):13–23, 2016. 2

[31] B. Settles. Active learning literature survey. `http://burrsettles.com/pub/settles.activelearning.pdf`, 2009. Accessed: 2025-06-14. 4

[32] K. Topping. Self and peer assessment in school and university: Reliability, validity and utility. In *Optimising new modes of assessment: In search of qualities and standards*, pp. 55–87. Springer, 2003. 2

[33] C. Valiente, J. Swanson, D. DeLay, A. M. Fraser, and J. H. Parker. Emotion-related socialization in the classroom: Considering the roles of teachers, peers, and the classroom context. *Developmental Psychology*, 56(3):578, 2020. 2

[34] 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*, 35:100434, 2022. 2

[35] W. Wagner, R. Göllner, S. Werth, T. Voss, B. Schmitz, and U. Trautwein. Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, 108(5):705, 2016. 2

[36] T. Xu, Q. Chen, Z. Zhang, B. Dong, H. Zhang, J. Bai, and Y. Zhou. Maximizing effectiveness of ai-generated instructors through human-like behavior and dynamic visual cues in instructional videos: Evidence from an eye-tracking study. *The Internet and Higher Education*, p. 101034, 2025. 2