Virtual reality in education: The impact of high-fidelity nonverbal cues on the learning experience

Vlasios Kasapakis [a,*], Elena Dzardanova [b], Androniki Agelada [a]

[a] Department of Cultural Technology and Communication, University of the Aegean, Greece
[b] Department of Systems Design and Engineering, University of the Aegean, Greece

ARTICLE INFO

ABSTRACT

Integrating variant Information and Communication Technologies (ICTs) into the learning process provides students and teachers alike with specialized tools which eliminate the distance between them and create a classroom-like experience. Virtual Reality (VR) is bound to not only match the qualities of interpersonal communication for distance-learning but reconfigure the learning process altogether, providing students and teachers with novel hyper-tools and methods for presentation and interaction. Virtual Reality Learning Environments (VRLEs) are already being designed, developed, and tested out as an educational tool. Among the less investigated aspects of VRLEs is the impact of avatars and characters Nonverbal Cues (NVCs) on the students' learning experience. This study presents the development of a prototype which uses off-the-shelf technologies commonly used in Social Virtual Reality (SVR) platforms to capture a real professor's body motion and gaze, along with his facial expressions, in real-time, during the delivery of a real lecture. The recorded data are later solved onto a high-fidelity avatar delivering the same lecture in a VRLE. A between-groups study including ninety-six (96) participants, all university students, revealed no correlation between the professor's avatar NVCs fidelity and perceived usability, realism, usefulness, and social presence, and no differences in knowledge acquisition as well.

1. Introduction

Virtual Learning Environments (VLEs), integrating several ICTs tools, are becoming common as they provide the option of distance learning (Al-Ajlan, 2012). With the use of web-conferencing services VLEs facilitate real-time lectures between remotely located students and teachers, supporting real-time voice communication, using voice over IP, live web-camera feed streaming, and multimedia content sharing (Gegenfurtner et al., 2020). This kind of technological services increase student engagement, highlighting VLEs as the best alternative to in-person learning (Kiss, 2012; Zhang et al., 2004). With VLE infrastructure and widespread usage already in place, the diffusion of immersive technologies, brought by the commercial re-invigoration of Virtual Reality (VR) as recently as seven years ago, is bound to provide its own niche for learning. Several studies are already reporting the positive impact that Virtual Reality Learning Environments (VRLEs) have on students' engagement and performance (Kavanagh et al., 2017).

The necessity of VRLEs became clear during the COVID-19 mandatory lockdowns when distance learning substituted classrooms and in-person learning worldwide (Singh et al., 2020), since the entire 2020–2021 academic year was in jeopardy for children in primary and secondary education, as well as university and college students (Mseleku, 2020). VRLEs however, were already being researched and proved effective in enhancing student engagement and enjoyment, along with easing knowledge acquisition through reduced cognitive effort (Kavanagh et al., 2017; Liagkou et al., 2019). One crucial aspect of the VR learning experience is the utilization of multi-user VRLEs as shared spaces or worlds wherein teachers and students have extensive control over 3D individual avatars. Studies on multi-user VRLEs and avatar representation are also reporting enhanced engagement and performance (O'Connor et al., 2018; Schild et al., 2018). Besides that, evolvement of immersive technologies is now looking toward incorporation of real-time body motion and gaze, as well as high-fidelity avatars by development of real-time facial cue capture (Chagué & Charbonnier, 2016; Roth et al., 2019). The technological capacities that allow Nonverbal Cues (NVCs) inclusion may significantly increase the quality of the learning

experience provided in VRLEs (Peeters, 2019; Theonas et al., 2008). Considering the rapid development of VR's comprising technologies and interfaces, VRLEs' potential as a distance-learning tool is clear. The aim of such alternatives ought to be not only a seamless learning experience but also a meaningful one. For that reason, several aspects pertaining to user experience, environment design, and system usability are still under thorough investigation regarding both the concerns and potential of VRLEs.

VRLEs interfaces (e.g., head-mounted displays, controllers, indoor motion tracking and so forth) have not been perfected yet, whereas amenities such as the ones provided by standalone head-mounted displays have only now become commercially available. However, since the vast majority of prospective users has yet to experience immersive technologies, let alone regularly, an adjustment period is expected, both for the obvious practicalities of using each system in a confined space, and for tailoring the educational methods and experience to the new medium and vice versa. These challenges, when combined with poor interaction and navigation design, can lead to VR learning and training experiences of reduced quality (Kavanagh et al., 2017; Sun & Gheisari, 2021). Furthermore, the degree of physical and cognitive task load increase during prolonged usage is not yet sufficiently evaluated to allow VRLEs to facilitate entire courses. Finally, the possibility of unexpected onset of negative side effects such as anxiety and cyber-sickness (Kamińska et al., 2019) is an important issue that has been extensively addressed, however, most experimental evaluations do not require prolonged concentration on an already cognitively strenuous activity such as learning, therefore actual population averages regarding duration tolerance and side effects may be undervalued.

The above issues are well known and at all times addressed by ongoing research as well as tackled through substantial investment in technological improvements. The growth rate of research and developer involvement is continuously increasing because, despite many shortcomings, there is also undeniable potential for VR and VRLEs in particular. NVCs integration, for instance, has been drawing increased research attention (Tanenbaum et al., 2020). That attention is mostly driven by the importance of NVCs in face-to-face interpersonal communication (Archer & Akert, 1977), which are considered equally crucial in virtual worlds (Roth et al., 2016; Smith & Neff, 2018). The addition of NVCs onto avatars and non-player characters contributes to their realism (Latoschik et al., 2017; Maloney et al., 2020) as they are an important communication tool in multi-user virtual worlds. However, detailed NVCs and/or transference of those from the real to virtual worlds is, in most cases, semi-automated (e.g., support of pre-programmed body gestures, automated facial expressions, lip syncing etc.). Such automated methods primarily add flavor to the remote virtual communication experience rather than provide users with a set of invaluable communication, thus affecting overall realism and social/interpersonal engagement (Tanenbaum et al., 2020). The latter is of great importance in VRLEs where social and behavioral cues are an integral part of the learning process (Ripka et al., 2020).

The present study explores the impact high-fidelity NVCs have on the learning experience of VRLE students. For this purpose, a VRLE prototype was developed, facilitating a real lecture by an actual professor, represented by an avatar. The professor's avatar featured several NVCs, specifically body motion, hand and finger gestures, facial expressions, and gaze motion, which were captured in real-time using a variety of technological solutions commonly used in Social Virtual Reality (SVR) platforms. The recorded data were applied to an avatar, which delivered a lecture to ninety-six (96) university students inside the VRLE prototype. Half of the students experienced the full spectrum of the professor's NVCs, while the other half experienced only his body motion and a subtle opening of the mouth when speaking. The students who took part in reviewing this virtual lecture completed a questionnaire providing valuable insights regarding the impact of the virtual professor's NVCs fidelity on their learning experience.

The remainder of this paper is structured as follows: Section 2 presents related work; Section 3 provides detailed information on the VRLE system architecture and design; Section 4 details the evaluation process and respecting results regarding the VRLE's usability, realism, usefulness, social presence, task load, and knowledge acquisition. Finally, Section 5 discusses and concludes this work, suggesting future research directions.

2. Related work

Professors use NVCs, such as facial expressions and body language, to convey important messages to students during real lectures, positively affecting their response and behavior (Feldman, 1986). Butt and Iqbal (2011) interviewed professors on the effectiveness of facial expressions during lectures. The study's results revealed professors consider facial expressions as an invaluable tool which helps maintain students' attention, increasing the quality and effectiveness of the learning experience. However, according to professors, certain facial expressions (e.g., expressions of anger) may hinder student attention. Neill (1989) examined students' response to slides depicting professors featuring different facial expressions and postures, demonstrating that smiling and frowning expressions had a positive and negative effect on students' learning experience, respectively. In contrast, the effect of professors' posture on students' learning experience was insignificant.

Previous studies also investigated the effects of NVCs on learning experiences delivered through conventional VLEs. Muilenburg and Berge (2005) reviewed the barriers of online learning to reveal that the lack of social cues during lectures delivered through VLEs led to communication uncertainties and feelings of isolation. Maimaiti et al. (2021) also observed the importance of NVCs in preventing student disengagement during lectures delivered through VLEs. Wang et al. (2019) compared student response on video lectures featuring different levels of professor expressiveness (heightened and conventional facial expressions). The study showed a positive impact of heightened facial expressions on the overall learning experience quality. Regarding knowledge acquisition, heightened facial expressions improved student's medium-term recall, but had no effect on student's short-term recall.

Theonas et al. (2008) developed and evaluated one of the earliest examples of a VRLE supporting NVCs. In this work, students attended pre-rendered virtual lectures, the virtual professor of which either featured or lacked facial expressions. The authors divided the virtual lectures based on text difficulty. The study's results show that students attending difficult-text lectures, the virtual professor of which featured facial expressions, performed 86% better in post-lecture tests than students attending the same lecture but with the virtual professor lacking facial expressions. The virtual professor's facial expressions did not have any effect on students' performance during easy-text lectures. Also, most of the students were more attentive during the lecture, the virtual professor of which featured facial expressions.

Another example of NVCs integration to VRLEs was introduced by Dharmawansa et al. (2015). Authors used a web camera to transfer students' eye blinking and head posture from the real world to an avatar inside a VRLE. According to the evaluation results of this study, both cues helped establish rich communication between students, enhancing the learning experience quality.

Zizza et al. (2018) developed and evaluated a multi-user VRLE using an HTC Vive Head-Mounted Display (HMD) along with its motion controllers. During this study, both professors and students controlled their own avatar, the body of which moved according to the real-word position of their head and hands capturing in this manner NVCs relating to body posture. Some students accessed the virtual world using non-immersive equipment, such as desktops or laptops. Students collaborated under the direction of the professor to solve several tasks in the virtual world (e.g., determine individual roles on a deserted island). The study's results show that both the immersive and non-immersive experiences of the VRLE were more immersive and enjoyable compared to conventional VLEs (iSocial).

Another example of a VRLE incorporating NVCs is presented by Ripka et al. (2020), where preservice teachers were immersed in a shared virtual space which simulated an auditorium and supported synchronous intercommunication. Each student accessed the virtual space using an Oculus Rift S HMD and was assigned a simple avatar consisting only of a head, torso, and hands which were disconnected from the rest of the avatar's body. During testing, teachers had to collaborate for the completion of certain tasks, such as arranging each planet of our planetary system as per their correct order. Evaluation results report enhanced attention to the subject at hand and overall fewer distractions in the VRLE compared to the physical world. The study states that lack of social and behavioral cues made communication over the collaborative tasks more challenging, even though participants reported experiencing a strong sense of community during their time in the VRLE.

Although there are studies that highlight the positive impact of NVCs integration in VRLEs, those are scarce. The VRLE prototype investigated in this study differentiates from other examples by introducing a high-fidelity avatar representing a real professor who had been previously captured delivering an actual lecture. The professor's avatar features multiple NVCs such as facial expressions, gaze, and full body motion, including finger motion, allowing for detailed hand gestures. This setup captures NVCs of higher fidelity compared to the examples found in the relating body of work. The aim of this study is a thorough investigation of NVCs effect on the learning experience to determine necessity for incorporating them, and to what extent, in similar commercial VRLEs developed in the future. Specifically, our setup takes under consideration the main drawbacks of the VRLEs described above and thus designs the VRLE presented here to evaluate, in correlation to the virtual professor's NVCs fidelity impact on the learning process, its overall usability, realism, usefulness, social presence, task load and knowledge acquisition. The hypotheses of this study are that students who experience a virtual professor featuring high-fidelity NVCs will (a) perceive the VRLE experience as more realistic, (b) will score higher the VRLE overall usefulness and its ability to draw their focus on the lecture, (c) and will experience higher levels of social presence.

3. System architecture

To develop the system presented in this work, we integrated and customized when required a variety of off-the-shelf technologies. Considering that VRLEs rely on real-time communication between professors and students, we opted to use technologies which can allow real-time communication between remote immersed users (Kasapakis et al., 2021). Thus, we excluded solutions for full-face motion capture as they cannot be used for real-time communication in SVR platforms where individuals' HMDs cover half of their face (Dagnes et al., 2019).

The core of the system comprises a VR Ready PC, a VIVE Pro Eye HMD, and two (2) base stations used to track its location in the real world. Five (5) VIVE Trackers were used to track the motion of the pelvis, hands, and feet of a real professor. Combining the VIVE Pro HMD and the VIVE Trackers with the Inverse Kinematics solution Final IK, allowed us to capture and transfer the full body motion of the real professor to his avatar in real-time. The overall development of the system was completed using Unity and SteamVR Software Development Kit (SDK), while the avatar representing the real professor in the virtual world was designed using the Reallusion Character Creator.

To track the professor's finger motion, we used ManusVR Gloves, ManusVR Apollo software, and ManusVR Unity SDK. Finger tracking, when using ManusVR Unity SDK, was not optimal, as several inconsistencies were documented. To accurately transfer finger rotation data from ManusVR Gloves to the professor's avatar, we authored additional code to calibrate each finger's rotation data using coefficient offsets. The capture of the professor's facial expressions required two separate solutions. The embedded eye tracker of HTC VIVE Pro Eye HMD was used to capture eye motion, therefore direction of gaze, which was then transferred to the professor's avatar via HTC VIVE Pro Eye SDK.

BinaryVR captured, through an infrared camera, motions of the mouth and lower jaw and transferred them onto the professor's avatar via an array of BlendShapes, which, as explained below, were also calibrated using coefficient offsets to match the professor's facial expressions. By this, aside from capturing any grimaces, the mouth movements of the avatar also matched and were accurately synchronized to the professor's speech, which was recorded during the lecture using the built-in HTC Vive Pro Eye HMD microphone. The completed system allowed us to record, using Unity Runtime Animation Recorder, the body and gaze motion, and facial expressions, of the real professor, export them in a Unity Animation File, and later transfer them onto his avatar on demand while integrating his voice too. The complete system architecture is presented in Fig. 1a, while the detailed system flow is presented in Fig. 1b.

The BinaryVR Unity's SDK includes an example of transferring facial expressions data onto the face of an avatar through BlendShapes. However, the model provided by BinaryVR is custom-made and unsuitable for solving tracked facial expressions onto other 3D human models. To address this issue, the BlendShape data provided by the BinaryVR SDK had to be adjusted accordingly to match those of the professor's avatar and thus, eliminate differences in the 3D model structure. A custom script helped create a float coefficient for each BlendShape. By manipulating each coefficient using a slider, while multiplying it with the corresponding BlendShape data provided by BinaryVR, we could calibrate any facial expressions offsets, matching the facial expressions of the professor in the real-world with the ones of his avatar in the virtual world (see Fig. 2a, Fig. 2b and 2c).

During the lecture, we recorded the body motion of the real professor using a digital camera, and his facial expressions using the raw feed of the BinaryVR infrared camera. We also recorded his gaze focus point in the virtual word, represented as a yellow sphere, using Vive Pro Eye SDK. We chose this solution as Vive Pro Eye SDK does not allow access to raw infrared camera feed because of privacy concerns (see Fig. 2d). We combined the recorded data into a single video file, maintaining the same frame per second rate between the digital camera and the screen recording of the virtual lecture. This allowed us to assess the accuracy between the real and the virtual world. We recorded the lecture five (5) times and after carefully reviewing the recorded data; we selected the most accurate recording to be evaluated by students.

4. User evaluation

4.1. Evaluation scenario

To evaluate our system's setup as a VRLE featuring high-fidelity NVCs, we integrated a virtual 3D auditorium in Unity. We also used Character Creator to design several non-player characters (NPCs) that helped simulate fellow students and a crowded classroom. We implemented these details (see Fig. 4) to provide the students who participated in the evaluation process with a classroom-like experience as close to reality as possible.

Prior to the evaluation process, we asked the real professor to prepare a short presentation based on an existing curriculum. The professor prepared a PowerPoint presentation on Virtual and Mixed Reality Technologies comprising several slides that included text, figures, and short clips/videos. We then integrated the PowerPoint file into the virtual classroom and provided the professor with two virtual buttons to help him navigate the presentation as he normally would in the real world. The equipment that helped capture his movements in real-time was appropriately calibrated so he could deliver the entirety of his lecture within the final VRLE at his own pace, engaging non-verbally with the virtual students (i.e., eye-contact, pointing) as he normally would in a real auditorium. The final motion capture was a recording of approximately fifteen (15) minutes in duration. The lecture duration was based on the average duration of past VRLE studies, which was about thirteen (13) minutes, delivering lectures within a certain time limit, featuring no
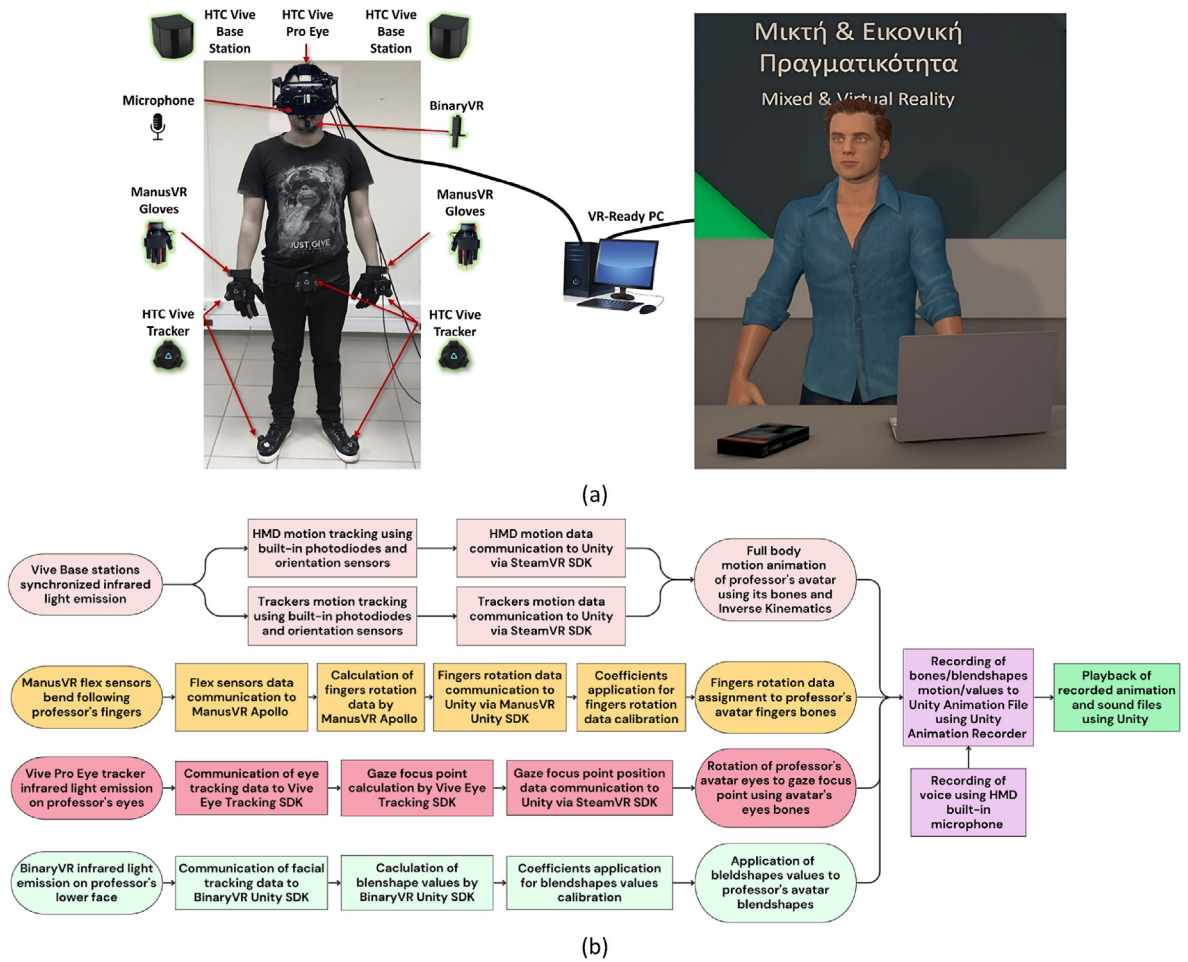
(a)

(b)

Fig. 1. (a) System architecture; (b) system Flowchart.
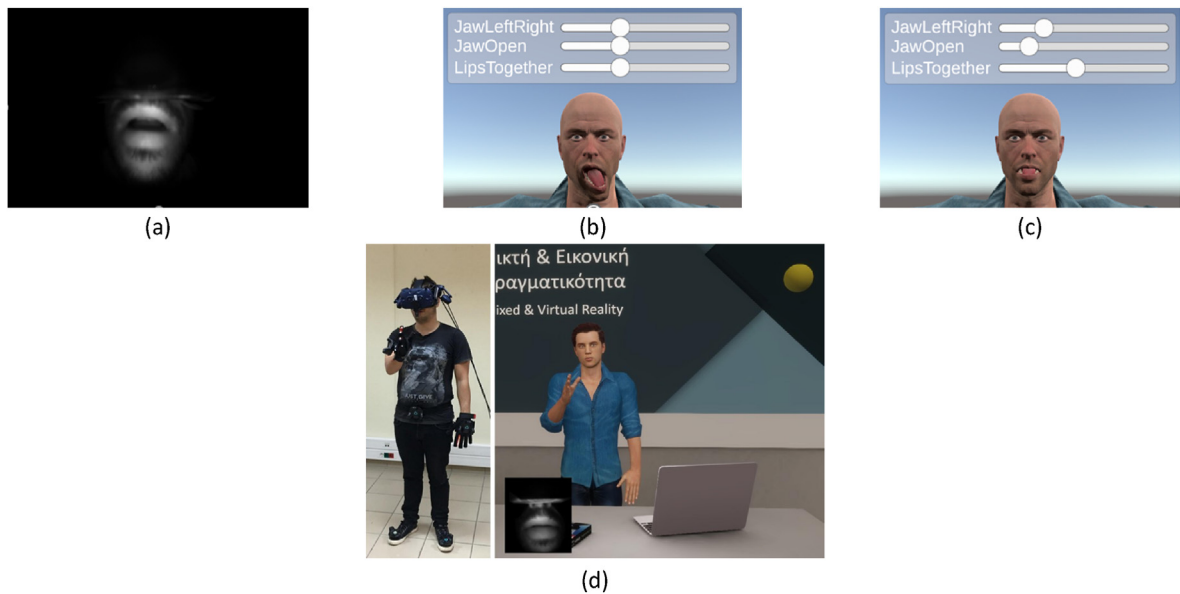

(a)   (b)   (c)

(d)

Fig. 2. (a) User facial expression in the real world as captured by the BinaryVR infrared camera; (b) Avatar facial expression prior to float coefficient application; (c) Avatar facial expression appropriately calibrated after float coefficient application; (d) Recorded data accuracy assessment.

interaction between professors and students (Hamilton et al., 2021). We also considered the increased effectiveness of focused and short-timed learning experiences delivered through VRLEs (Villena-Taranilla et al.,

2022).

As already mentioned, this was a between-groups study, therefore both groups experience the same lecture except for the variable under

evaluation, which are the professor's NVCs. Group A experienced the full spectrum of NVCs as those were previously captured, consisting of full body motion, hand and finger gestures, facial expressions of the lower face, and gaze. Whereas for Group B, the original capture was tampered with, maintaining only full body motion, excluding finger gestures, and applying an automated opening of the mouth in sync with the professor's speech. The rest of the facial expression of the virtual professor of Group B was set to the default neutral expression of Character Creator Studio. We did not provide any specific directions to the professor regarding facial expressions or gestures to be performed during recordings. Considering that a real professor delivered the lecture, and not an actor, this would interfere with the naturalness of his motion and affect the communication of NVCs. A comparison of video recordings between Group A and Group B lectures revealed that apart from increased gaze fidelity, Group A students experienced one hundred and fifty-four (154) detailed finger gestures (e.g., pointing, waving, and finger counting) performed by the professor, and one hundred and eight (108) distinct facial expressions recorded when the professor smiled, smirked, or emphasized on certain words to increase student engagement. Thus, the professor employed several NVCs to obtain and maintain student attention during the lecture, creating a difference in the experience between Group A and Group B students, as shown in Fig. 3.

A total of ninety-six (N96) participants, all university students between 18 and 24 years old, participated in the evaluation process. We equally divided students between Group A (N48) and Group B (N48), and they experienced the VRLE lecture (see Fig. 4a) mostly in groups of seven (7). A researcher assisted students with the mounting, comfort, and lens adjustments of HTC Vive Pro HMDs that were used to access the VRLE and was there at all times monitoring the sessions. Each group was given approximately 1 min to observe the classroom and the fellow student NPCs that were populating it (see Fig. 4b). A variety of "idle sitting" animations, exported from Mixamo, were applied to the NPCs, while their heads were adjusted to face the professor. The virtual fellow student NPCs featured eye-blinking in Group A, a feature that was removed in Group B. Finally, an "empty room" ambient sound looped for the duration of the virtual lecture to help with immersion and block any external sounds. The virtual lecture application scene faded out and in, revealing the virtual professor who greeted the students and delivered the lecture (see Fig. 4c).

Upon completion of the sessions in the VRLE, we asked students to complete a questionnaire evaluating their experience. The factors examined to determine NVCs effect were usability, realism, usefulness,

social presence, task load, and knowledge acquisition. Questions regarding usability, realism and usefulness were custom to the aim and objectives of this study. Questions regarding social presence were largely based on Koskela et al. (2018), while those regarding task-load assessment were based on the Nasa Task Load Index (Hart, 2006).

The questionnaire also included questions evaluating participant previous experience with (a) digital applications (e.g., games) that incorporate characters and avatars; (b) Immersive Virtual Reality systems and applications; and (c) familiarity with the lecture's content (VR and MR technologies). Those questions were scaled from 1 (no experience at all) to 10 (substantial experience). As shown in Fig. 5, previous experience differences between Group A and Group B students were very small (see Fig. 5 Q1, Fig. 5 Q2, and Fig. 5 Q3). This was also verified by a paired inferential statistic ($p = 0.1$ in all three questions) showing that the levels of previous experience in both groups are not different.

4.2. Evaluation results

Apart from the questions regarding previous experience with relevant applications, systems, educational content, and task load, which were ten-point, all other questions used a five-point Likert-Scale (Strongly Disagree–Strongly Agree). To properly present the Likert-scale questions results, we used both descriptive and inferential statistics (Allen & Seaman, 2007). First, we treat them as ordinal data to get a general impression of each question. Then, we extract the average score of each question per evaluated aspect (e.g., usability, realism, etc.) and per participant, and present the spread of the resulting interval data. Finally, we execute paired statistical hypothesis tests for each set of interval data to reveal any statistical significance.
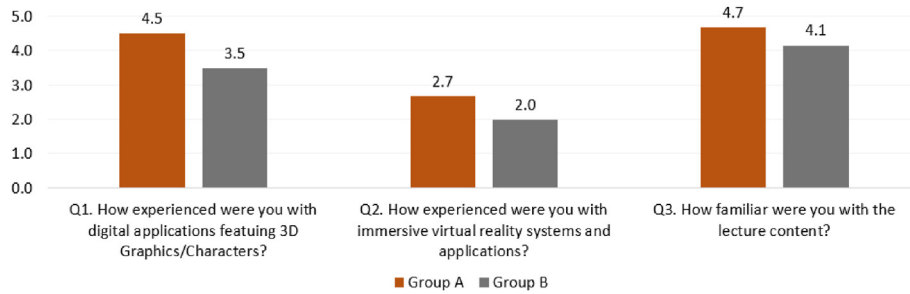
4.3. Usability

Prior to presenting insights about the impact of the virtual professor's NVCs fidelity on students' learning experience, it is important to discuss the overall perceived usability of the VRLE to ensure that there was no technical issue which could affect the evaluation results. As shown in Fig. 6, almost all students agree that the VRLE application ran smoothly (see Fig. 6 Q1), while there were no latencies regarding the virtual professors' overall motion and body movements (see Fig. 6 Q2). Almost all students could clearly observe the virtual professor (see Fig. 6 Q3) and hear his voice in high quality (see Fig. 6 Q4), which was also properly synced with his motion and body movements (see Fig. 6 Q5).
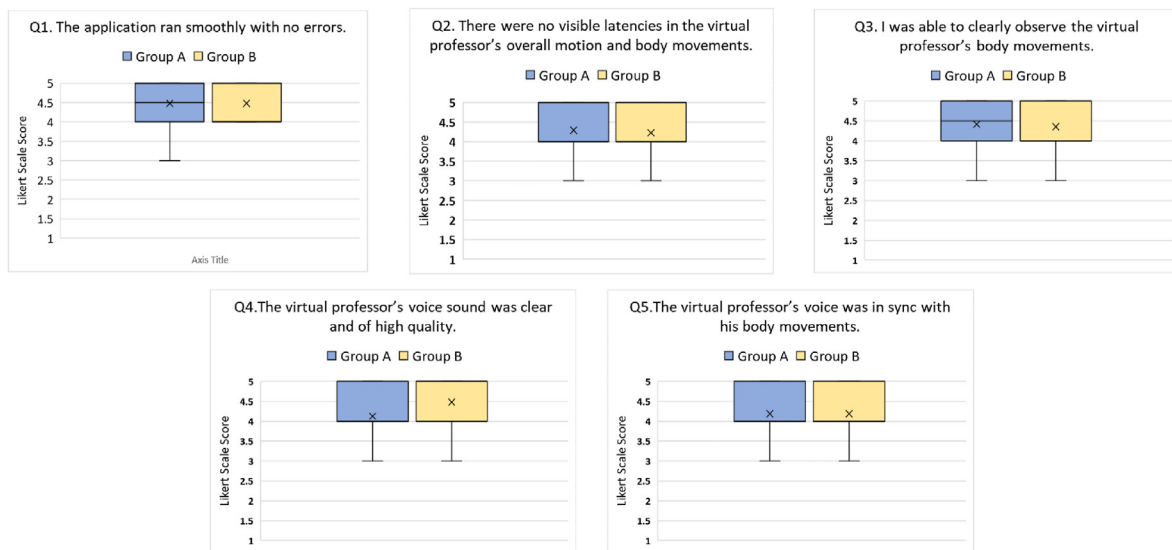

Fig. 3. Virtual professor featuring reduced (left) and full (right) NVCs during the same animation frame.
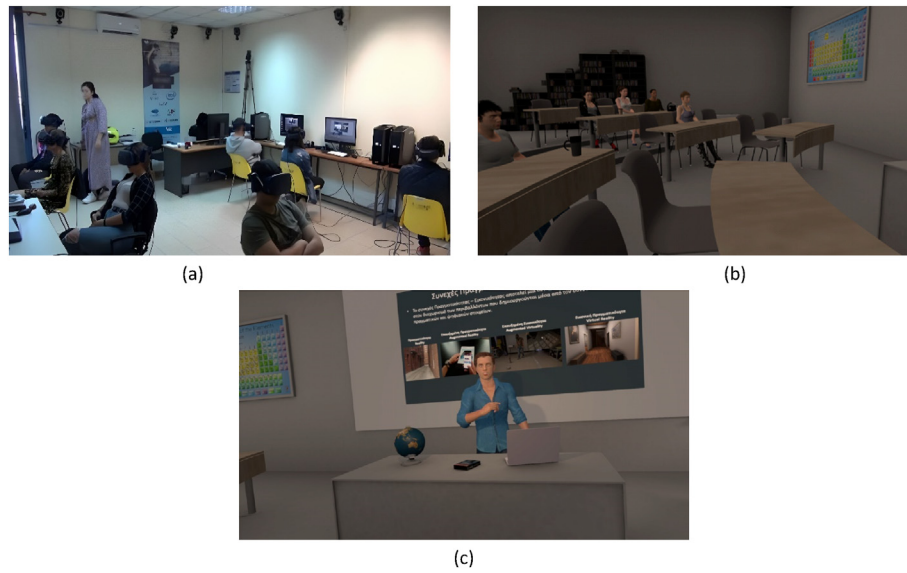
Fig. 4. (a) Students participating in the evaluation process of the virtual lecture; (b) Student view while observing virtual fellow student NPCs; (c) Student view of the professor during the virtual lecture.

Fig. 5. Previous experience questions average scores.

Fig. 6. Perceived usability between Group A and Group B.

Group A students answered an additional question regarding the virtual professor's facial expressions and gaze direction clarity. As shown in Fig. 7, Group A students could clearly observe both of the virtual professor's aforementioned NVCs.

Fig. 8 presents the interval data of the usability questions (we excluded the additional question answered by Group A students to ensure data validity). As shown, there are no great differences between Group A and Group B regarding perceived usability. A paired statistical hypothesis
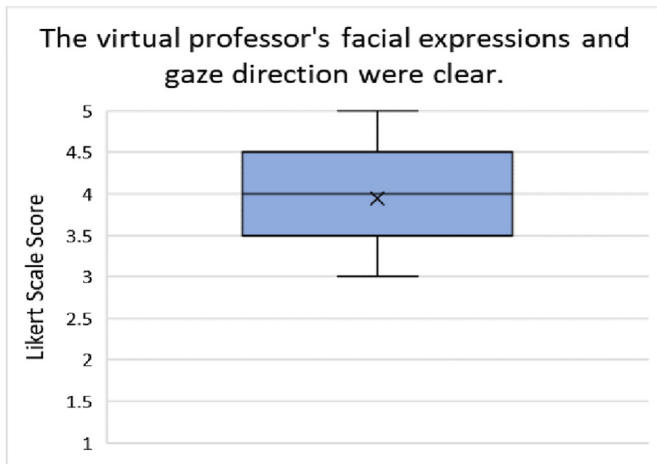
Fig. 7. Group A usability question regarding the virtual professor's facial expressions and gaze motion.

test revealed no statistical significance ($p = 0.3$) between the results of the two groups as well.

This section's results verify students could clearly listen and observe the virtual professor during the lecture, while his NVCs were clear and in sync with his voice.

4.4. Realism

Regarding realism, Fig. 9 shows no differences between the two groups. Both groups' students agreed that the overall virtual lecture was realistic (see Fig. 9 Q1), but to a reduced extent compared to a real lecture with a real professor (see Fig. 9 Q3) and real fellow students (see Fig. 9 Q4). The overall realism of the virtual professor (see Fig. 9 Q2), and the realism of his body movements (see Fig. 9 Q5) scored the same adequate levels of realism in both groups.

Group A students, who experienced the full spectrum of the virtual professor's NVCs, answered two additional questions regarding perceived realism. Most of the students perceived both virtual professor's facial expressions and gaze as realistic (see Fig. 10 Q1). Almost all students agreed that the detailed virtual professor's NVCs increased the virtual lecture's level of realism (see Fig. 10 Q2).

Fig. 11 presents the interval data of the questions regarding perceived realism (we excluded the two additional questions answered by Group A students to ensure data validity). The results show that there are no significant differences regarding overall perceived realism. A paired statistical hypothesis test showed no significant difference ($p = 0.2$) between Group A and Group B as well.

In conclusion, the evaluation results regarding realism, both per question and interval, show that Group B students perceived the VRLE as realistic as students of Group A. This is an interesting result that refutes the hypothesis that Group A students, who experienced a virtual professor featuring high-fidelity NVCs, would render the overall experience as more realistic.
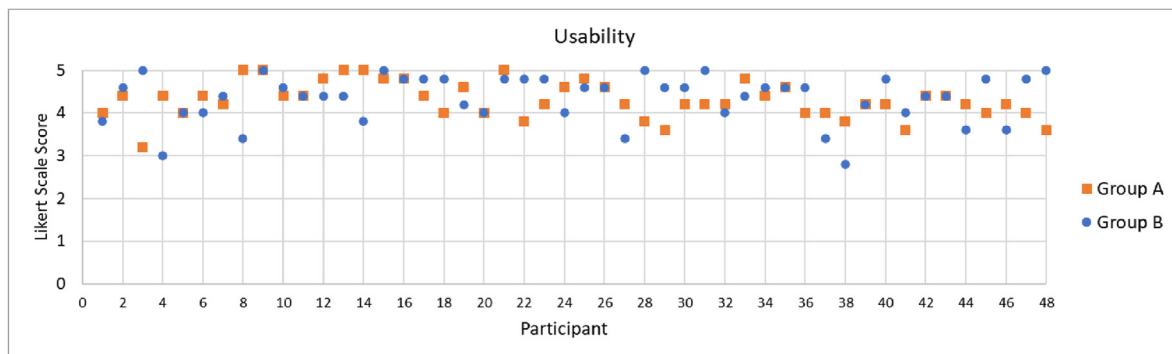

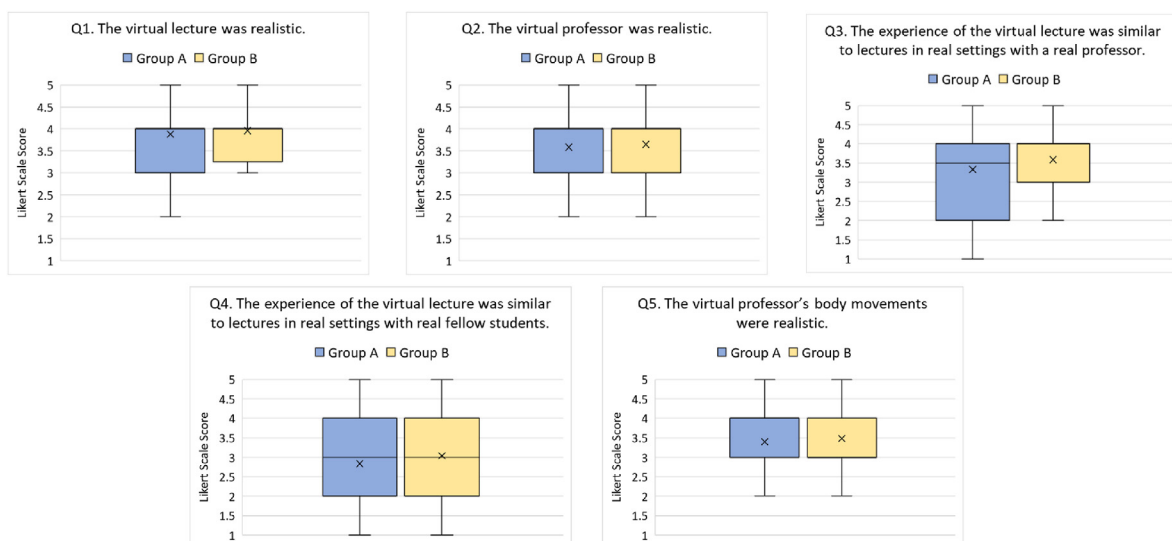**Fig. 8.** Usability interval scores between Group A and Group B.

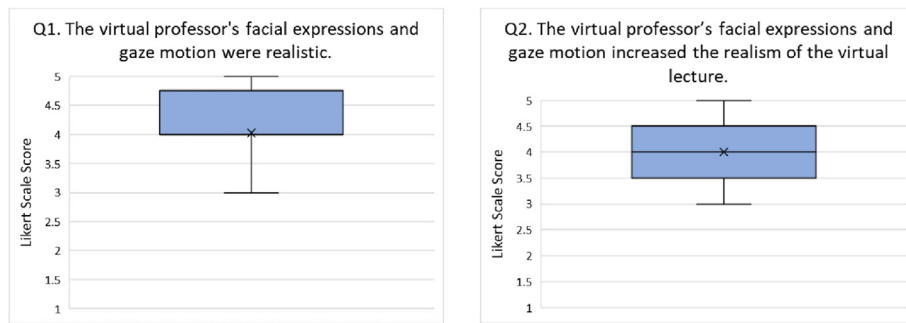**Fig. 9.** Perceived realism between Group A and Group B.

Fig. 10. Group A questions regarding the virtual professor's facial expressions and gaze realism.
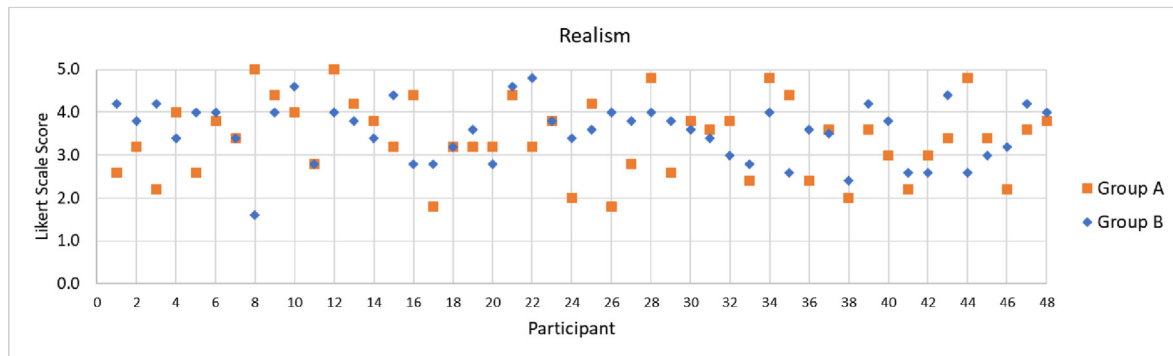

Fig. 11. Realism interval scores between Group A and Group B.

4.5. Usefulness

The virtual professor's usefulness, as shown in Fig. 12, was positively evaluated by almost all students in both groups. First, students regarded the lecture to be of high quality (see Fig. 12 Q3), further increased by the professor's body motion (see Fig. 12 Q8). Students acknowledged the professor assisted them in understanding the lecture's content (see Fig. 12 Q1), while he successfully guided and kept their attention (see Fig. 12 Q2, Fig. 12 Q4, and Fig. 12 Q6). Finally, most of the students stated they understood the lecture's content (see Fig. 12 Q5), while the overall body motion of the virtual professor added towards that statement (see Fig. 12 Q7).

Group A students were asked to complete three (3) additional questions regarding the usefulness of the virtual professor's facial expressions and gaze motion. As shown in Fig. 13, almost all students agree that the detailed NVCs of the virtual professor assisted keeping their attention and helped them understand the lecture's content. Including those NVCs increased the overall quality of the lecture.

As verified by the interval scores (see Fig. 14), Group A and Group B expressed the same opinions regarding the overall usefulness of the virtual professor showing that incorporation of high-fidelity NVCs made no significant difference (we excluded the additional questions answered by Group A to ensure data validity). This was also verified by a paired statistical hypothesis test ($p = 0.4$) which showed no statistical significance between the two groups.

4.6. Social presence

As Fig. 15 shows, students of both groups felt that during the lecture the virtual professor was present in the same space as them (see Fig. 15 Q1), and vice versa, they felt themselves to be in the same space as the virtual professor (see Fig. 15 Q4). The effect of social presence triggered by virtual fellow student NPCs was lower compared to the virtual professor, but still above average (see Fig. 15 Q2 and see Fig. 15 Q5). The same stands for the professor's perceived awareness of the virtual fellow

student NPCs (see Fig. 15 Q7). However, students of Group A felt less acknowledged by the virtual professor during the lecture (see Fig. 15 Q6), while students in both groups were not sure when the virtual professor was focusing directly on them (see Fig. 15 Q3).

Overall, the evaluation results of this section reveal that the high-fidelity NVCs had no significant impact on students' perceived social presence. This is also verified by the interval results presented in Fig. 16. There were no statistically significant differences between the two groups, as shown by a paired statistical hypothesis test ($p = 0.1$).

4.7. Task load

To evaluate the perceived task load for attending the virtual lecture we used, and appropriately adjusted, the Nasa Task Load Index. The scale of the questions was set from 1 to 10. We found no significant differences regarding cognitive (see Fig. 17 Q1) and physical effort (see Fig. 17 Q2) required to attend the virtual lecture. As stated by students of both groups, the pace of the lecture was balanced (see Fig. 17 Q3) and easy to attend to (see Fig. 17 Q5). Finally, most students felt that they successfully attended the lecture (see Fig. 17 Q4), while there were only a few cases where students felt insecure, discouraged, or irritated during the lecture (see Fig. 17 Q6).

4.8. Knowledge acquisition

To test knowledge acquisition, the real professor designed and provided us with a multiple-choice quiz featuring ten (10) questions that covered the key points of the virtual lecture's content. Each multiple-choice question had three (3) options to ensure the validity of the results (Vyas & Supe, 2008). We asked students to complete multiple-choice quiz after completing the evaluation questionnaire. Group A and Group B participant scores were about the same. A paired statistical hypothesis test revealed no significant difference ($p = 0.1$) on the impact of the virtual professor's NVCs on the knowledge acquisition of students.

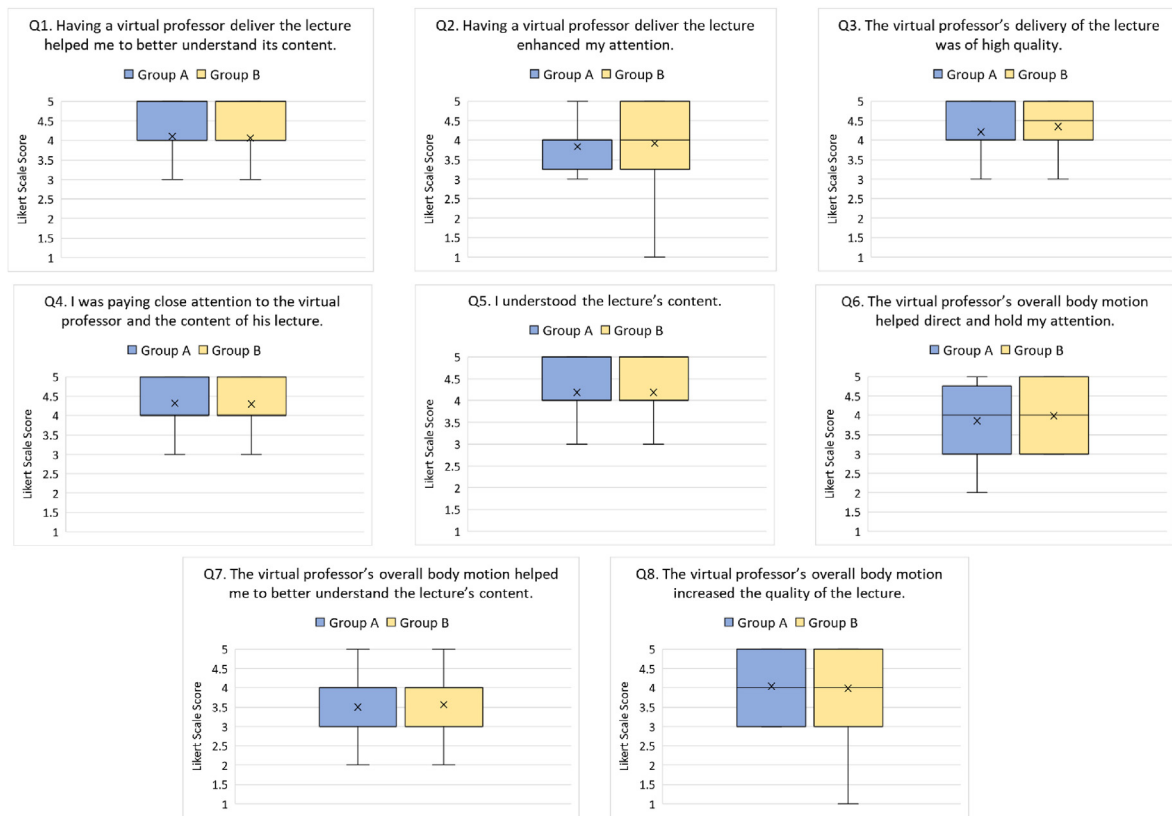To appropriately investigate knowledge acquisition, we excluded

Fig. 12. Perceived usefulness of the virtual professor between Group A and Group B.
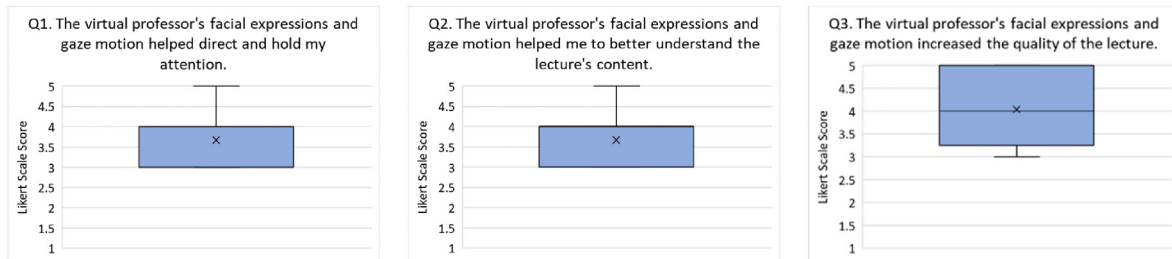

Fig. 13. Group A questions regarding virtual professor's facial expressions and gaze motion usefulness.
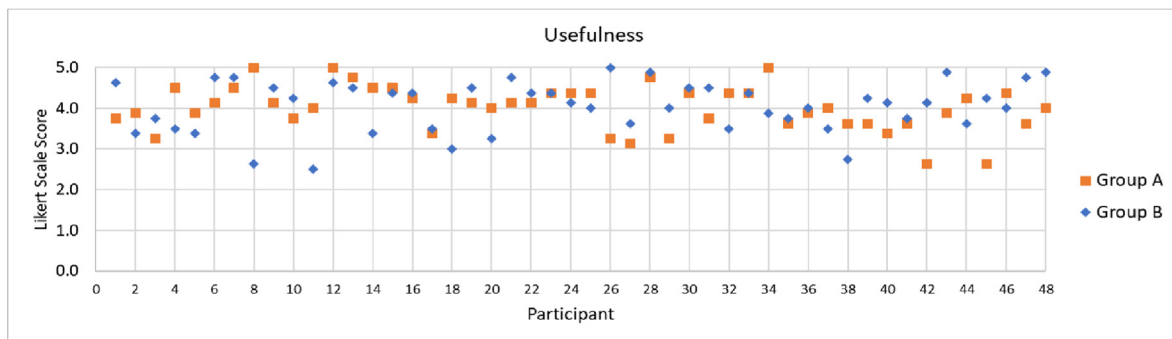

Fig. 14. Usefulness interval scores between Group A and Group B.

students with self-reported previous knowledge on the lecture's content. Seventeen (17) students of Group A and twenty-three (23) students of Group B had no previous knowledge of the lecture's content. Still, no significant differences were found between the two groups (see Fig. 18).
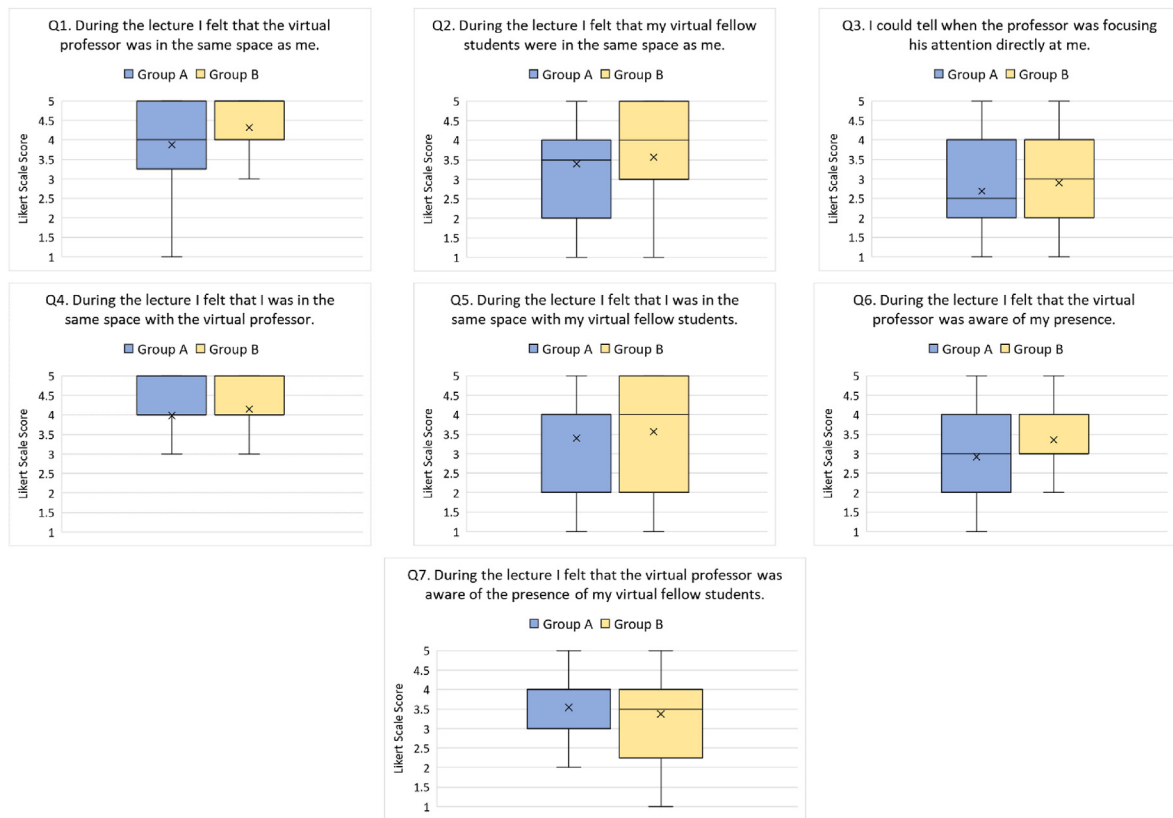
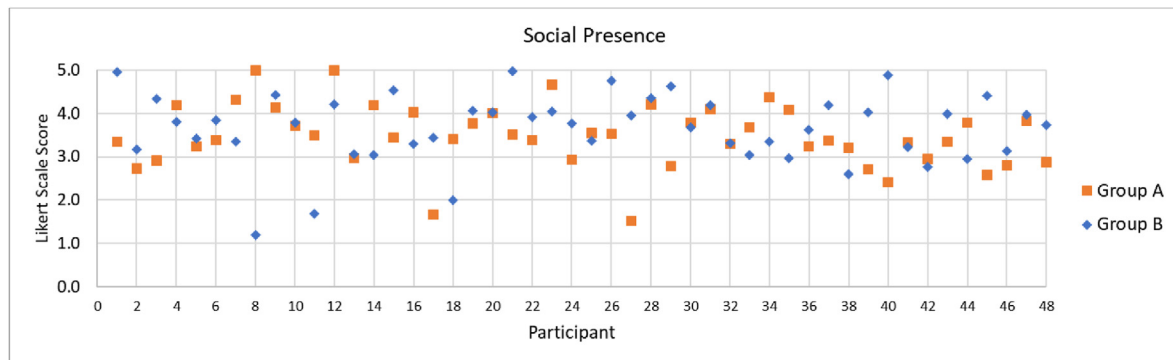**Fig. 15.** Perceived social presence between group A and group B.

Fig. 16. Interval Social Presence scores between Group A and Group B.
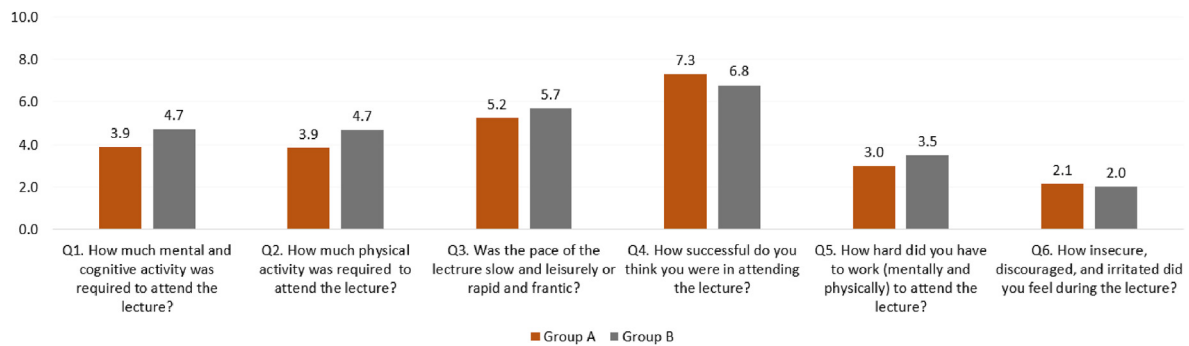

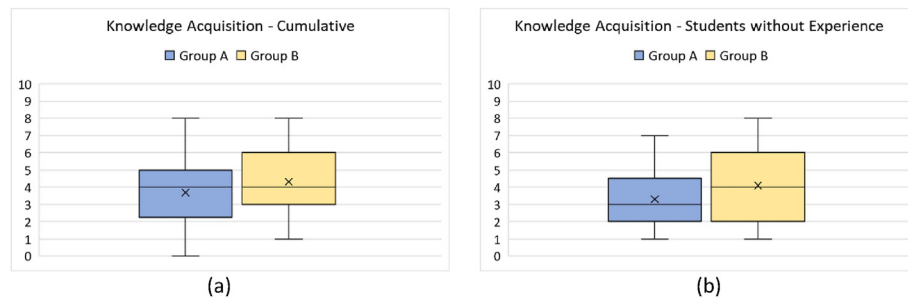**Fig. 17.** Perceived task-load between Group A and Group B.

Fig. 18. Multiple-choice quiz results between Group A and Group B.

An ANOVA analysis of variance revealed no statistically significant differences as well ($p = 0.1$).

5. Discussion

The evaluation results of the study confirmed the proposed system's usability as no technical or other issues arose and students of both groups were able to smoothly observe the virtual lecture and its content, and clearly see and evaluate the NVCs presented to them. Despite the setup's overall successful function, the between-groups study refutes all three hypotheses according to which Group A evaluation would score higher across all factors examined as those are listed above. Statistical analysis shows no significant difference between the two groups and no differences in knowledge acquisition. All statistical tests ran indicate that, in this particular experimental setup, NVCs as the differentiating variable had no significant effect compared to the version of the setup that lacked them. These results do not necessarily discredit previous studies that document the effect of NVCs on participant experience, as there are several additional points to consider.

The first point relates to the overall fidelity and high quality of VRLEs, meaning that overall improved 3D graphics, smoothness of applications, and high-quality VR equipment and immersive technologies, might by themselves negate any necessity for developmental additions relating to certain NVCs, at least in this kind of setups where there is no one-on-one intercommunication between parties. NVCs are not limited to an interlocutor's body posture and facial expressions, but also to the social context of the interaction (Dzardanova et al., 2022). In this case, the 3D environment, seems to easily and readily affirm for students that this is a learning experience, by design of the auditorium, usage of typical props such as projection of a presentation, and appropriate positioning of parties that indicates social hierarchy and conventional social order between them. By this, this study's evaluation results indicate that such environment-related features might suffice in conveying the nature of the experience.

Another point to consider is a more consistent real-life application of VRLEs between real professors, their own students, and the actual curriculum and syllabus that they need to go through. Real-life applications would, first of all, vary greatly in duration, which in turn could affect parameters such as attention, task load and/or social presence. Real-life relations also might come with varying shifts in social dynamics as indicated by computer-mediation studies that far extend VR settings and provide great insight regarding issues of anonymity or compliance to authority (Dzardanova et al., 2022). NVCs are detrimental when considering these points, thus further experimentation is required that tests out VRLEs, perhaps in less controlled settings.

Also, our setup lacked any social interaction and/or interactive options. For example, in a real setting students might be required to verbally contribute during a lecture, deliver a presentation of their own, they would also talk amongst each other, take notes, use different media, aside from the fact that one of the greatest incentives for utilizing VRLEs is their ability to simulate complex concepts as exemplified in the study by Ripka et al. (2020) where participants move around our planetary system

(Ripka et al., 2020). Therefore, whether a VRLE application opts for high-fidelity representation is more a matter of how developmentally challenging and sustainable such additions are, rather than how necessary they might be to ensure the quality of the learning experience.

Regarding knowledge acquisition, there is a key factor which may have led to similar student performance in both groups. The multiple-choice test examined the short-term recall of students. As observed in another study conducted using video lectures, the short-term recall of students was not affected by the professor's facial expressions and body language fidelity (Wang et al., 2019). This is a rather interesting result as that study, which video lectures features similar duration as ours, showed similar students' short-term recall scores in both heightened and conventional professor expressivity, as observed in our work too, thus showing that students short-term recall may not be affected by professors' NVCs in both VLEs and VRLEs, when those provide short learning experiences.

Another factor which may have led to similar levels of student attitude towards the virtual professor's NVCs, and knowledge acquisition in both Group A and Group B is exposure duration. Commonly, related research efforts, along with the one presented here, evaluate short passive learning experiences in VRLEs, which can be rather effective (Villena-Taranilla et al., 2022). However, as Hamilton et al. (2021) mention in their literature review, students may not fully grasp the benefits of learning through VRLEs in short-termed experiences, as they may need more time to acclimate to the environment. Considering the subtle role NVCs have in interpersonal communication (Phutela, 2015), their impact may require a prolonged exposure of students to VRLE-enabled learning experiences.

The most common issues preventing the adoption of VRLEs in education are the cost of the required technological infrastructure and expert personnel, along with the limited technological orientation of educators (Marks & Thomas, 2022). According to the findings of our study, compared to low-fidelity NVCs, integrating high-fidelity NVCs in VRLEs that offer passive, brief, and to the point learning experiences did not improve the overall learning experience quality and student knowledge acquisition. Implementing low-fidelity or automated NVCs in short learning experiences provided through VRLEs significantly decreases their design, development, and equipment costs, easing their adoption to real-world education. Finally, revealing that high-fidelity NVCs do not affect short term learning experiences in VRLEs, eases the access of educators to simplified platforms for VRLEs design and development (e.g., Mozilla Hubs, Frame etc.). Such platforms commonly support semi-automated NVCs out of the box, which can be integrated to the learning process taking place inside the virtual world, providing students with rich learning experiences.

6. Conclusions and future work

The aim of the present study is the evaluation of the importance of NVCs during virtual lectures facilitated in VRLEs. To this end we designed a prototype VRLE simulating an auditorium and populated with fellow student NPCs to resemble a classroom-like experience as much as

possible. A real professor was previously recorded delivering a real lecture whilst his body, hands, fingers, lower face facial expressions, and gaze were motion captured with appropriate off-the-shelf technologies, customized when necessary.

Ninety-six (96) students in total participated in the system setup evaluation and filled out extensive questionnaires to determine how impactful the virtual professor's NVCs were on perceived usability, realism, usefulness, and social presence during the learning experience. Student knowledge acquisition was also tested out. Students were equally divided between two groups, with Group A experiencing the full spectrum of NVCs as those were previously recorded in real time, and Group B only experiencing the professor's full body motion and subtle opening of the mouth to be in sync with his verbal delivery.

Our study is largely conclusive that NVCs fidelity does not affect the most important factors pertaining to education in VRLEs wherein students are passive, and the learning experience is rather short and to the point. Notwithstanding, this still leaves the great variety of intercommunication and interactive, as well as collaborative contexts as mentioned above, wherein NVCs might still be critical and therefore future evaluations are required. Examples of such contexts can vary from VRLEs integrating increased social interaction, and used to deliver any learning material allowing educators and students to interact in real-time with each other using high-fidelity avatars, or ones developed for specific purposes requiring increased NVCs fidelity, such as sign language learning (Dzardanova et al., 2022), non-verbal communication training for individuals with autism (Georgescu et al., 2014), and medical students training on patients assessment based on nonverbal communication (Deladisma et al., 2007). Finally, future studies should also investigate factors such as prolonged exposure and medium/long-term recall of students to further examine the impact of NVCs on the learning experience delivered through VRLEs.

Statements on open data and ethics

The participants were protected by hiding their personal information in this study. They were voluntary and they knew that they could withdraw from the experiment at any time. The data can be provided upon requests by sending e-mails to the corresponding author.

Declarations of competing interest

None.

Acknowledgement

References

Al-Ajlan, A. (2012). *A comparative study between e-learning features. Methodologies, Tools and New Developments for E-Learning* (pp. 191–214).

Allen, E., & Seaman, C. (2007). Likert scales and data analyses. *Quality Progress, 40*(7), 64–65.

Archer, D., & Akert, R. (1977). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology, 35*(6), 443.

Butt, M. N., & Iqbal, M. (2011). Teachers' perception regarding facial expressions as an effective teaching tool. *Contemporary Issues In Education Research, 4*(2), 11–14.

Chagué, S., & Charbonnier, C. (2016). Real virtuality: A multi-user immersive platform connecting real and virtual worlds. In *Proceedings of the virtual reality international conference* (pp. 1–3).

Dagnes, N., Marcolin, F., Vezzetti, E., Sarhan, F.-R., Dakpé, S., Marin, F., Nonis, F., & Mansour, K. B. (2019). Optimal marker set assessment for motion capture of 3D mimic facial movements. *Journal of Biomechanics, 93*, 86–93.

Deladisma, A. M., Cohen, M., Stevens, A., Wagner, P., Lok, B., Bernard, T., Oxendine, C., Schumacher, L., Johnsen, K., & Dickerson, R. (2007). Do medical students respond empathetically to a virtual patient? *The American Journal of Surgery, 193*(6), 756–760.

Dharmawansa, A., Fukumura, Y., Marasinghe, A., & Madhuwanthi, R. (2015). Introducing and evaluating the behavior of non-verbal features in the virtual learning. *International Education Studies, 8*(6), 82–94.

Dzardanova, E., Kasapakis, V., Vosinakis, S., & Psarrou, K. (2022). Sign Language in immersive VR: Design, development, and evaluation of a testbed prototype. In *Proceedings of the 28th ACM symposium on virtual reality software and technology* (pp. 1–2).

Feldman, R. (1986). Introduction: The present and promise of the social psychology of education. *Social Psychology of Education: Current Research and Theory, 5*.

Gegenfurtner, A., Zitt, A., & Ebner, C. (2020). Evaluating webinar-based training: A mixed methods study of trainee reactions toward digital web conferencing. *International Journal of Training and Development, 24*(1), 5–21.

Georgescu, A. L., Kuzmanovic, B., Roth, D., Bente, G., & Vogeley, K. (2014). The use of virtual characters to assess and train non-verbal communication in high-functioning autism. *Frontiers in Human Neuroscience, 8*, 807.

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). 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, 8*(1), 1–32.

Hart, S. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting* (pp. 904–908).

Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R. E., Avots, E., Helmi, A., Ozcinar, C., & Anbarjafari, G. (2019). Virtual reality and its applications in education: Survey. *Information, 10*(10), 318.

Kasapakis, V., Dzardanova, E., Nikolakopoulou, V., Vosinakis, S., Xenakis, I., & Gavalas, D. (2021). Social virtual reality: Implementing non-verbal cues in remote synchronous communication. In *International conference on virtual reality and mixed reality* (pp. 152–157).

Kavanagh, S., Luxton-Reilly, A., Wuensche, B., & Plimmer, B. (2017). A systematic review of virtual reality in education. *Themes in Science and Technology Education, 10*(2), 85–119.

Kiss, G. (2012). Comparison of traditional and web-based education-case study "BigBlueButton". In *Proceedings of the international symposium on information technologies in medicine and education* (pp. 224–227).

Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., & Botsch, M. (2017). The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM symposium on virtual reality software and technology* (pp. 1–10).

Liagkou, V., Salmas, D., & Stylios, C. (2019). Realizing virtual reality learning environment for industry 4.0. *Procedia Cirp, 79*, 712–717.

Maimaiti, G., Jia, C., & Hew, K. F. (2021). Student disengagement in web-based videoconferencing supported online learning: An activity theory perspective. *Interactive Learning Environments*, 1–20.

Maloney, D., Freeman, G., & Wohn, D. Y. (2020). Talking without a voice" understanding non-verbal communication in social virtual reality. In *Proceedings of the ACM on human-computer interaction* (pp. 1–25).

Marks, B., & Thomas, J. (2022). Adoption of virtual reality technology in higher education: An evaluation of five teaching semesters in a purpose-designed laboratory. *Education and Information Technologies, 27*(1), 1287–1305.

Mseleku, Z. (2020). A literature review of E-learning and E-teaching in the era of Covid-19 pandemic. *International Journal of Innovative Science and Research Technology, 57*(52), 588–597.

Muilenburg, L., & Berge, Z. (2005). Student barriers to online learning: A factor analytic study. *Distance Education, 26*(1), 29–48.

Neill, S. (1989). The effects of facial expression and posture on children's reported responses to teacher nonverbal communication. *British Educational Research Journal, 15*(2), 195–204.

O'Connor, M., Deeks, H. M., Dawn, E., Metatla, O., Roudaut, A., Sutton, M., Thomas, L. M., Glowacki, B. R., Sage, R., & Tew, P. (2018). Sampling molecular conformations and dynamics in a multiuser virtual reality framework. *Science Advances, 4*(6), eaat2731.

Peeters, D. (2019). Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review, 26*(3), 894–900.

Phutela, D. (2015). The importance of non-verbal communication. *IUP Journal of Soft Skills, 9*(4), 43.

Ripka, G., Grafe, S., & Latoschik, M. E. (2020). Preservice teachers' encounter with social VR–exploring virtual teaching and learning processes in initial teacher education. In *Proceedings of the SITE interactive conference* (pp. 549–562).

Roth, D., Bente, G., Kullmann, P., Mal, D., Purps, C. F., Vogeley, K., & Latoschik, M. E. (2019). Technologies for social augmentations in user-embodied virtual reality. In *Proceedings of the 25th ACM symposium on virtual reality software and technology* (pp. 1–12).

Roth, D., Lugrin, J.-L., Galakhov, D., Hofmann, A., Bente, G., Latoschik, M. E., & Fuhrmann, A. (2016). Avatar realism and social interaction quality in virtual reality. In *Proceedings on virtual reality and 3D user interfaces* (pp. 277–278).

Schild, J., Lerner, D., Misztal, S., & Luiz, T. (2018). EPICSAVE—enhancing vocational training for paramedics with multi-user virtual reality. In *Proceedings of the 6th international conference on serious games and applications for health* (pp. 1–8).

Singh, R. P., Javaid, M., Kataria, R., Tyagi, M., Haleem, A., & Suman, R. (2020). Significant applications of virtual reality for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research Reviews, 14*(4), 661–664.

Smith, H. J., & Neff, M. (2018). Communication behavior in embodied virtual reality. In *Proceedings of the CHI conference on human factors in computing systems* (pp. 1–12).

Sun, Y., & Gheisari, M. (2021). Potentials of virtual social spaces for construction education. *EPiC Series in Built Environment, 2*, 469–477.

Tanenbaum, T. J., Hartoonian, N., & Bryan, J. (2020). How do I make this thing smile?" An inventory of expressive nonverbal communication in commercial social virtual

reality platforms. In *Proceedings of the CHI conference on human factors in computing systems* (pp. 1–13).

Theonas, G., Hobbs, D., & Rigas, D. (2008). Employing virtual lecturers' facial expressions in virtual educational environments. *International Journal of Virtual Reality, 7*(1), 31–44.

Villena-Taranilla, R., Tirado-Olivares, S., Cózar-Gutiérrez, R., & González-Calero, J. A. (2022). *Effects of virtual reality on learning outcomes in K-6 education: A meta-analysis.* Educational Research Review, Article 100434.

Wang, Y., Liu, Q., Chen, W., Wang, Q., & Stein, D. (2019). Effects of instructor's facial expressions on students' learning with video lectures. *British Journal of Educational Technology, 50*(3), 1381–1395.

Zhang, D., Zhao, J. L., Zhou, L., & Nunamaker, J. F., Jr. (2004). Can e-learning replace classroom learning? *Communications of the ACM, 47*(5), 75–79.

Zizza, C., Starr, A., Hudson, D., Nuguri, S. S., Calyam, P., & He, Z. (2018). Towards a social virtual reality learning environment in high fidelity. In *Proceedings of the 15th annual consumer communications & networking conference (CCNC)* (pp. 1–4).