Effect of Look-Alike Avatars on Students' Perceptions of Teaching Effectiveness

Kwame Agyemang Baffour *
Graduate Center, City University of New York

Oyewole Oyekoya†
Hunter College, City University of New York

Figure 1: Avatar Representations (close-up and VR condition): (1) Look-Alike Avatar; (2) Video-Avatar; (3) Stick-Avatar

ABSTRACT

This paper presents a study investigating the influence of look-alike avatars on students' perceptions of teaching effectiveness in three-dimensional virtual and augmented reality environments. This study investigated three avatar representations: i) look-alike avatar of the instructor; ii) stick avatar; and iii) video recording of the instructor. Eighteen participants were asked to rank the three avatar representations, as well as the immersion experience of the teaching simulation on the virtual and augmented reality displays. The result of this study suggests that look-alike avatars can be used to represent an instructor in virtual environments.

Index Terms: Computing methodologies—Virtual reality—;——Computing methodologies—Perception

1 INTRODUCTION

Over the last few years, technology has become increasingly available that allows people to make avatars that look like and behave like the user (look-alike avatars). The creation of photorealistic self avatars with full face and body animations is important for virtual reality applications that aim for perception and action to replicate real world experience [2].

Prior studies have mainly focused on self-evaluation of this photo-realistic self avatars that look like the user. Less emphasis have been put on external person's evaluation, which is perhaps more objective. As such, this paper presents a study investigating the influence of look-alike avatars on students' perception of teaching effectiveness in immersive telepresence environments. To investigate this, we developed a VR and AR application where an instructor, whom the participants are familiar with (participants have taken instructor's class in the past), delivers the same lesson as a look-alike avatar, stick avatar or video avatar. Subsequently, we conducted subjective evaluations of the participants to compare how they rank the avatar representations of the instructor to the stick avatar and video recording in both VR and AR environments.

2 METHOD

18 unpaid participants (9 males and 9 females) were recruited for the study. All participants had normal or corrected-to-normal vision and no color blindness. Ages ranged from 18 to 34, with fifteen participants falling within the 18-24 age range. All participants had prior knowledge or use of virtual or augmented reality except one participant. The pre-experiment questionnaire also asked participants to rank their knowledge of Coronavirus disease (COVID-19) on a Likert scale of 0 (not much) to 5 (a lot), the median score was 4 with 17 participants ranking their knowledge at 3 or above.

The experiment used a 2-factor repeated measures (within-subjects) design. The independent variables were (i) avatar representation (stick-avatar, look-alike avatar and video-avatar) and (ii) display (VR and AR). Using a counterbalanced measures design, the avatar representations and displays were mixed to reduce any confounding influence of the order and sequence effects such as learning or fatigue. The look-alike avatar of the instructor was generated by scanning the head of the instructor using the 3dMDhead System (https://3dmd.com/) while the lower body was scanned and generated with a Structure Sensor (https://structure.io). The generated 3D models were merged and automatically rigged using Adobe Mixamo (https://mixamo.com). For the video avatar representation, a video recording was displayed on a 2D plane in the virtual environment. The video was a one minute recording of the instructor delivering a lesson on COVID-19. The voice of the instructor was also used for the look-alike and stick avatars in both the VR and AR conditions. The VR and and marker-based AR applications were developed in Unity3D using Google VR and Vuforia SDKs respectively. The smartphone used for the VR and AR applications was a Samsung Galaxy A32 with a screen size of 6.5 inches, refresh rate of 60-90Hz and a resolution of 720 x 1600 pixels.

Participants were asked to fill out a pre-experiment questionnaire about gender, age range, knowledge of VR, AR and COVID-19.Participants then watched the VR and AR simulations of the lessons being delivered by the three avatar representations, in turn. After the simulations, participants were asked to fill out a shortened form of the validated Immersion Experience Questionnaire [3]. Fourteen Immersion Experience questions (Questions 13 - 26) were selected and reworded to rate the overall experience of the displays on a 6-point Likert scale ranging from 0 (Very low/Not at all) to 5 (A lot/Very much so). The preference questionnaire was designed based on the teaching evaluation questions used by instructor's home institution (Hunter College). Twelve preference questions were used in selecting preferred avatar representation. Participants were asked to rate which avatar: (i) effectively maintained their interest and attention; (ii) did they prefer the most; (iii) did they feel most comfortable with; (iv) made them feel the strongest sense of closeness to their usual classroom instructor; (v) had the most influence on their understanding of the subject matter; (vi) makes clear, practical demonstrations; (vii) encourages independent and critical thought; (viii) makes good use of examples and illustrations to clarify concepts; (ix) interprets difficult and abstract ideas; (x) effectively communicates his/her knowledge of the subject to them; (xi) challenges them intellectually; and (xii) has increased their interest in the subject.

*e-mail: kagyemang@gradcenter.cuny.edu
†e-mail: oyewole.oyekoya@hunter.cuny.edu

3 RESULTS

Two dependent variables were measured: (i) avatar preference and (ii) immersion experience. For the avatar preference survey, an aggregate preference measure for each avatar representation was computed for each participant by counting the proportion of times, over the 12 questions, that a participant chose each method. Since the measure is a frequency/count, the data did not follow a normal distribution. The residual errors did not satisfy the tests of normality with a Shapiro-Wilks test, and the hypothesis of normality was rejected ($p < 0.05$) for all conditions. Hence, non-parametric Pearson's chi-square test was performed to assess the frequencies of participants' choices for the most preferred avatar representation across the two displays. There was a statistically significant association between the avatar representation and display ($\chi^2(2) = 29.274, p < 0.001$) i.e. the choices selected for the most preferred avatar representation significantly differed for both VR and AR. Bonferroni-adjusted comparisons of proportions showed that there was a statistically significant proportion of participants who preferred the look-alike avatar ($N = 125$) over the video avatar ($N = 59$) in the VR ($p < .05$) condition. Bonferroni-adjusted comparisons of proportions also showed that there was a statistically significant proportion of participants who preferred the video avatar ($N = 109$) over the look-alike avatar ($N = 72$) in the AR ($p < .05$) condition. The proportion of participants who preferred the stick avatar was not statistically significant by display ($p > 0.5$).

In our study, three dimensions of immersion (transportation, challenge and emotional involvement) were extracted from the study by Jennett et al. [3]. Responses to negatively scored items (questions 6 and 8) were first inverted. Cronbach's alpha ($\alpha = 0.856$) showed that the ratings have relatively high internal consistency. The responses to all fourteen questions are summed to give a single measure of immersion experience for each participant. The immersion scores were entered into a one-way repeated measures Analysis of Variance (ANOVA) with one factor of display. As there are only two levels of repeated measures, Mauchly's test of sphericity indicated that the assumption of sphericity had not been violated ($\chi^2 = 0.00$). The residual errors satisfied the tests of normality with a Shapiro-Wilks test, and the hypothesis of normality was not rejected ($p > 0.05$) for all display conditions. ANOVA results revealed that there was no statistically significant difference between displays ($F_{(1,17)} = 2.026, p > 0.05$) with similar immersion scores for the VR ($Mean = 41$) and AR ($Mean = 39$) displays. In other words, immersion scores for the VR and AR displays were rated similarly.

4 DISCUSSIONS

Results of the analysis showed that participants preferred the look-alike avatar in VR settings and video avatar in AR settings. As expected, the stick avatar was the least preferred avatar across all displays. Participants' ranking of their sense of immersion in the VR condition did not significantly differ from the AR condition. This suggests that participants perceived the immersion of the VR and AR environments similarly. Despite the similarity, participants significantly preferred the look-alike avatar in the VR setting, as opposed to the video avatar in the AR setting. This indicates that look-alike avatars are a viable avatar representation for instructors in teaching and demonstrating core concepts to students in VR environments. The preference for the video avatar in the AR setting suggests that participants' subjective evaluation of virtual avatars in real world is more critical, with higher expectations of the perceived realism of the avatars. This study suggests that the realistic match of the video to the real world plays a major factor in how participants perceive the video avatar. Also, it was easier to watch the video-avatar in the AR environment since the background of the instructor in the high-definition recording was constant and participants did not have to experience the changing contrasts of the background of the real environment against the virtual look-alike and stick avatars. This suggests that a 3D-video avatar (e.g. point-cloud rendering) may have yielded different results. Pakanen et al. [4] suggested that a hologram full body avatar might work best in AR settings due to its see-through appearance. These studies suggest that any perceived uncanniness of the look-alike avatar is more apparent in AR environments, as participants were more critical of its imperfections in the real world AR setting.

Based on written feedback, the stick avatar was generally regarded as creepy by most participants. According to some participants, the look-alike avatar showed a lot of potential to impact future educational VR/AR applications and was the most preferred avatar in VR. The video avatar however was the most preferred avatar in AR because it represents a more natural and familiar mode of instruction to participants.

This study was designed to enable the researchers to understand how participants who were familiar with the instructor's teaching would perceive the look-alike avatar of the instructor. Familiarity plays a role in the uncanny valley effect [1] and a highly personalized look-alike avatar of the instructor (voice and facial features) may explain why participants prefer the look-alike avatar in the VR setting. These results must be treated with caution, as it is not clear from this study whether the participants would have a similar judgment if the instructor was unfamiliar. However, the use case is viable, as students were forced to move from in-person to remote instruction at the onset of the COVID-19 pandemic. This results do suggest that a photorealistic full-body avatar in VR setting may be perceived by students as a viable alternative for remote instruction.

5 CONCLUSIONS AND FUTURE WORK

The aim of this study was to explore the effect of look-alike avatars on students' perception of teaching effectiveness in AR and VR settings. This was done by having a look-alike, stick and video avatar of an instructor that the participants were familiar with, deliver a lesson in VR and AR environments. The paper showed that look-alike avatars are preferred for representing the instructor in VR environments. Participants also have higher expectations of look-alike avatars in AR environments and clearly noticed imperfections, thereby preferring the 2D-video avatars in AR settings.

This study did not investigate any effects on learning but rather on the subjective evaluation of the avatar representations in AR and VR. Future work will focus on using advanced tools to develop more realistic look-alike avatars with full facial and body movements. Attention measures such as head gaze, target tracking and number of head fixations will also be recorded and analyzed in future work.

ACKNOWLEDGMENTS

This work was supported by a grant from PSC-CUNY (Award No. TRADB-51-248).

REFERENCES

[1] A. W. de Borst and B. de Gelder. Is it the real deal? perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in Psychology*, 6, 2015.

[2] A. Frampton-Clerk and O. Oyekoya. Investigating the perceived realism of the other user's look-alike avatars. In *Proc. of the 28th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–2. ACM, 2022.

[3] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66(9):641–661, 2008.

[4] M. Pakanen, P. Alavesa, N. van Berkel, T. Koskela, and T. Ojala. "nice to see you virtually": Thoughtful design and evaluation of virtual avatar of the other user in ar and vr based telexistence systems. *Entertainment Computing*, 40:100457, 2022.