Study of Cross-Reality Interfaces with Avatar Redirection to Improve Desktop Presentations To Headset-Immersed Audiences

Jason W. Woodworth* Christoph W. Borst†

University of Louisiana at Lafayette

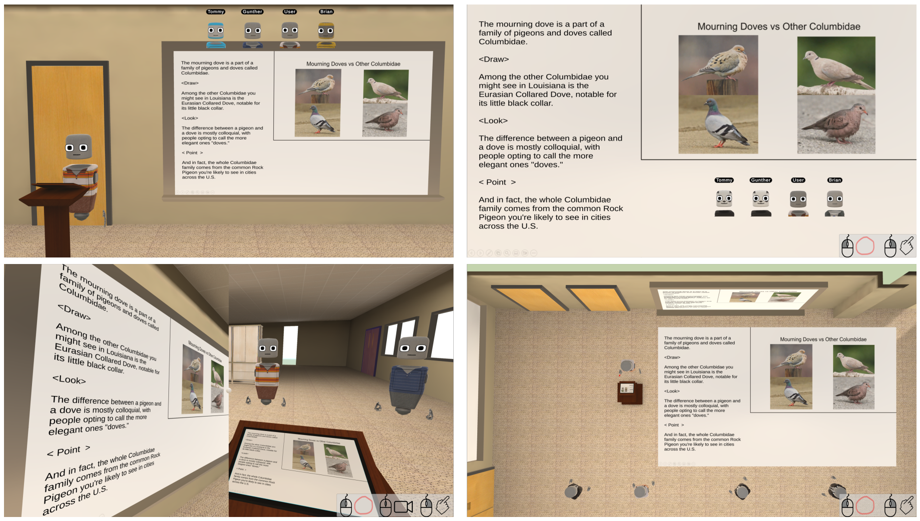

Figure 1: Top left: SDesk interface, showing the presenter and board from an audience perspective and using desktop-based eye tracking to control the presenter's avatar's gaze. Top right: PPT, like conventional desktop presenters, adapted to cross-reality use. Bottom left: Conventional Desktop, resembling common first-person VR-on-desktop view. Bottom right: TopDown, a bird's-eye view of the environment allowing the presenter to see all audience members in context of the virtual world.

Abstract

Recent work on VR meeting spaces for presentations suggests that some presenters prefer to use desktop interfaces for comfort or other factors. We designed and compared cross-reality desktop interfaces that allow a desktop user, with a 3D avatar, to present lecture-style content to a headset-immersed audience. In contrast to previous work, we now evaluate the presenter's experience with different interfaces, rather than the audience experience. Our SDesk interface, with an audience-type viewpoint and eye-tracked gaze gestures, is preferred over a standard VR-simulating desktop interface.

Index Terms: Human-centered computing—Interaction design—Empirical studies in interaction design;

1 Introduction

Increased availability of VR meeting spaces has been driven by a combination of the rising availability of VR devices, remote classes and meetings, and the spread of work-from-home jobs. However, the requirement of VR devices comes with noteworthy drawbacks. While VR devices offer precise spatial movement, finely tracked avatars, and immersive displays, desktop interfaces offer better ergonomics and easier access to peripherals like keyboards, in addition to currently being more ubiquitous. These differences have led to the increased consideration of cross-reality (CR) interfaces that allow users of different device types to meet and interact in a shared virtual space. Woodworth et al. introduced a set of CR desktop interfaces designed to animate a presenter's meeting space avatar in a way that would mimic human motion more naturally than conventional desktop interfaces [2]. However, the interfaces were only validated from the audience's perspective. We now summarize a study focusing on the presenter's experience.

2 Cross-Reality Interface Design

The four studied interfaces (seen in Fig. 1) differ in what view the user is given of a classroom and how certain actions are performed. To animate a presenter avatar for cross-reality use, each interface works with the Body Mechanics system from [2]. Desktop inputs, such as mouse pointing or eye tracking, are redirected onto avatar motion. All interfaces allow the user to see a presentation board and some representation of students. If the interface does not naturally show the full student avatar, they are represented by "cutouts" placed as UI elements on the screen, allowing the presenter to monitor them for raised hands and status indicators. When the board is in view, each interface allows the user to direct their avatar to point at / draw on the board by right / left clicking on the desired area.

PowerPoint-Style (PPT) imitates features of standard desktop presenter views. The screen shows slide content in one corner and notes on what to say and do on the side. Like a video-chat tool's webcam views, student cutouts are presented below slide imagery. Hovering the mouse cursor over the face of an intended student causes the teacher's avatar to look at that student. Note this view does not let the presenter see their own avatar.

TopDown (TD) presents a top-down classroom view of the teacher, students, and board. As the original board is unusable due to the perspective, an interactable duplicate board is drawn to the center

*e-mail: jason.woodworth1@louisiana.edu
†e-mail: cwborst@gmail.com

This is an author-formatted version. Original publication: J. W. Woodworth and C. W. Borst, "Study of Cross-Reality Interfaces with Avatar Redirection to Improve Desktop Presentations to Headset-Immersed Audiences," 2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Orlando, FL, USA, 2024, pp. 1082-1083, doi: 10.1109/VRW62533.2024.00337.

right. The teacher avatar looks at a student when the teacher hovers the mouse over the student avatar. When a student raises a hand, an additional hand icon is drawn atop the avatar.

SDesk (SD) is our sensors-based interface from [2]. Imagery is shown from a student-like perspective to give a sense of being in the classroom. The teacher's avatar is rendered left of the board so the teacher can see how it appears to students. Visual cutouts of student avatars appear in a row above. An eye tracker detects when the teacher looks at a cutout, and maps this to avatar gaze at a student.

Conventional Desktop (CD) shows a first-person view like common interfaces for desktop CR, e.g., Mozilla Hubs. The user can rotate the camera by click-drag with a mouse. The teacher's avatar looks in the camera direction, so users rotate the camera to face students. A non-interactive duplicate of the board is placed on the podium to give users some idea of content while looking at students.

3 Usability Study

Forty-seven subjects (30 male, 17 female, aged 18 to 38, mean age 21.1) from a computer science department used the interfaces to give 4 approximately 2-minute presentations (one presentation per interface type, in random order) to a group of 4 automated students. Presentations were kept simple and only intended to give context to the teaching activity. Text and instructions were given to subjects, located next to slide imagery. The desktop had a 32" 4K monitor with a Tobii Eye Tracker 5.

Before using each interface type, subjects previewed a physical printout of the presentation slides, then completed a brief tutorial on the upcoming interface, then performed the presentation. Subjects read presentation text aloud to simulate teaching. For task instructions (pointing at or drawing on the board, or looking at a student) they had to complete the task before continuing.

After each interface, subjects answered 10 questions from the System Usability Scale [1]. At the end, they ranked each interface by: "Which was your **Favorite** to use?", "Which was the **Easiest** to use?", and "Which would you rather have a **Teacher** use?".

4 Results

We analyzed results with Friedman and Holm-corrected Wilcoxon tests. Fig. 2 summarizes aggregated SUS score. There was a significant moderate effect of interface on SUS ($\chi^2(3) = 40.075, p < .001, W = .304$). Followup tests showed differences between most pairs, giving a partial ranking. CD (median SUS of 77.5) had lower score than SD ($p = .005$), TD ($p < .001$), and PPT ($p < .001$). SD (median SUS of 87.5) had lower score than TD ($p = .006$) and PPT ($p < .001$). TD (92.5) and PPT (95) were not shown different ($p = .062$), but the borderline result may suggest a likely trend. We establish a possible best-to-worst ranking of PPT > TD > SD > CD. We also note that all four interfaces achieved a median SUS score above 68, implying they all have above average usability [1].

Fig. 2 also shows rankings. A Friedman test revealed a difference in Favorite ($\chi^2(3) = 13.555, p = .004, W = .103$). Followup comparisons only revealed that SD (median rank of 2) was ranked better than CD (median rank of 3, $p = .008$).

There was a moderate significant difference in Ease rankings ($\chi^2(3) = 64.609, p < .001, W = .489$). Followups revealed differences between most pairs. CD (median rank of 4) was considered harder to use than SD, TD, and PPT ($p < .001$ for all). PPT (median rank 1) was considered easier than SD ($p = .009$) and TD ($p = .009$). No difference was shown between SD (median 2) and TD (median 2.5, $p = .731$). We infer a partial ranking of PPT > SD / TD > CD.

A Friedman test showed a difference in rankings of preference for teacher use ($\chi^2(3) = 20.836, p < .001, W = .158$). Followups showed SD (median rank of 2) was ranked better than CD (median 3, $p = .003$) and TD (median 3, $p = .004$). PPT (median 2) was also ranked better than CD ($p = .026$).

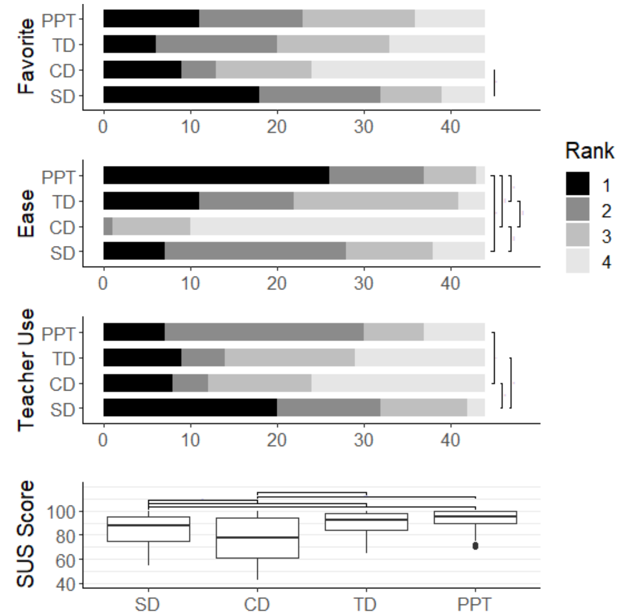

Figure 2: Top: Ranking for Favorite, Easiest, and preferred Teacher use. Bar color indicates how many subjects gave the indicated ranking. Bottom: Boxplots showing SUS score for each interface. Connecting brackets imply significant differences between pairs.

5 Discussion

For favorability, SD received the highest number of top rankings, while CD received the highest number of bottom rankings. Subjects who ranked SD highly noted that they liked the perspective, the element layout, and the novelty of eye tracking. The 9 subjects who ranked CD as their favorite all noted they liked the feeling of immersion, even if it was more tedious to use. A ranking for TD and PPT are less clear, but we consider it most likely that they rank somewhere between SD and CD.

Rankings and SUS scores both show CD at the bottom. While subjects who favored it noted they liked the immersive feel of the first-person perspective, a majority said it was unnecessarily complicated compared to the others. While SD was not ranked easiest to use, subjects appeared to enjoy using it, stating that they liked the perspective, element layout, and novelty of eye tracking. Subjects also thought teachers would like the layout and the ease of eye tracking if it could be calibrated well enough.

TD was often ranked somewhere in the middle. While some appreciated the ability to see the full classroom, most found it unnecessary and did not like the layout, with four saying specifically that it felt distracting or unnatural. The main appeal of PPT was its ease of use and familiarity, but many subjects ranked it under SD for Teacher use because of the appeal of eye tracking, claiming they thought it would feel more natural for teachers.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1815976.

References

[1] J. Brooke. Sus: A quick and dirty usability scale. *Usability evaluation in industry*, 189, 11 1996.

[2] J. W. Woodworth, D. Broussard, and C. W. Borst. Redirecting desktop interface input to animate cross-reality avatars. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 843–851, 2022. doi: 10.1109/VR51125.2022.00106