Perception and Proxemics with Virtual Humans on Transparent Display Installations in Augmented Reality

Juanita Benjamin* University of Central Florida
Gerd Bruder† University of Central Florida
Carsten Neumann‡ University of Central Florida
Dirk Reiners§ University of Central Florida
Carolina Cruz-Neira¶ University of Central Florida
Gregory F. Welch‖ University of Central Florida

Abstract

It is not uncommon for science fiction movies to portray futuristic user interfaces that can only be realized decades later with state-of-the-art technology. In this work, we present a prototypical augmented reality (AR) installation that was inspired by the movie *The Time Machine (2002)*. It consists of a transparent screen that acts as a window through which users can see the stereoscopic projection of a three-dimensional virtual human (VH). However, there are some key differences between the vision of this technology and the way VHs on these displays are actually perceived. In particular, the *additive light model* of these displays causes darker VHs to appear more transparent, while light in the physical environment further increases transparency, which may affect the way VHs are perceived, to what degree they are trusted, and the distances one maintains from them in a spatial setting. In this paper, we present a user study in which we investigate how transparency in the scope of transparent AR screens affects the perception of a VH's appearance, social presence with the VH, and the social space around users as defined by proxemics theory. Our results indicate that appearances are comparatively robust to transparency, while social presence improves in darker physical environments, and proxemic distances to the VH largely depend on one's distance from the screen but are not noticeably affected by transparency. Overall, our results suggest that such transparent AR screens can be an effective technology for facilitating social interactions between users and VHs in a shared physical space.

Keywords: Augmented reality, virtual humans, appearance, proxemics, social presence.

1 Introduction

Transparent screens are an emerging technology that is being considered for natural user interfaces in a wide range of fields. In particular, movies such as the *Iron Man* trilogy, *Avatar (2009)*, and *Minority Report (2002)* have a long history of leveraging transparent screens so that the audience is able to see both the actors and the on-screen computer interfaces they are interacting with. Other movies such as *The Time Machine (2002)* turned this method around to show a virtual human (VH) through the transparent screen as if it were a real person standing behind a glass divider. While the vision of this technology in the Science Fiction movies is intriguing, the actual technology we have available to build these setups remains comparatively under-explored in the virtual reality (VR) and augmented reality (AR) research community [25, 41]. Advances in transparent screens and projection foils made it possible to build installations that are similar to the envisioned setups [28, 35, 38]. However, there are some critical differences, and we do not yet know how the display technology may affect interactions between users and VHs.

Transparent screens, similar to optical see-through (OST) head-mounted displays (HMDs) like the Microsoft HoloLens, follow a different light model than traditional TVs and computer screens. The *additive light model* of transparent screens means that light reaching the user's eyes consists of a blend of (a) light coming from the display and (b) light coming from the physical environment behind the screen [14, 15]. The practical effect is that these displays can only "add light" but not "take light away" from the user's view of the physical environment. It means that the color black can not be induced on the display as it denotes the absence of light. Rendering black pixels on the display turns them transparent, through which the user will see the color of the physical background behind the display. Previous research showed that this can cause issues with traditional user interfaces, suggesting that certain color schemes have benefits over others, such as "dark mode" graphics (e.g., characterized by light foreground text on a dark transparent background) result in details being generally more legible than traditional "light mode" graphics (e.g., dark foreground text on a light background) [13, 26]. However, recent studies also revealed that these issues transfer over to human representations as skin color schemes are affected by the display technology as well, meaning that humans with darker skin colors, hair, or clothes will appear more transparent than humans with lighter colors [36], which can be exacerbated in situations with high amounts of light in the physical environment [11, 13].

The issues related to darker colors and transparency are not simply inconveniences. For example, Peck et al. found that the more transparent the skin of a VH (e.g., due to darker skin), the more the VH can be dehumanized [36]. It stands to reason that a VH displayed on a transparent screen could very well have the same issues (darker skin appearing transparent), and those issues could also give rise to dehumanization. Further, such perceived dehumanization could be erroneously attributed to a negative demeanor or intent on the part of the VH, which could in turn affect a person's proxemics in the presence of the VH. Just as transparency issues are not simply inconveniences, altered proxemics can be related to trust [6, 40] and could impact the effectiveness of the VH.

In this paper, we present a user study (N=22), where we evaluated a transparent AR screen prototype with different amounts of background light (physical environment) and foreground light (VHs with different skin colors), and different distances at which the VHs appeared relative to the screen and relative to the participants. With our study we addressed the following two research questions related to this promising display technology:

RQ1 What are the effects of AR display technology on the perceived *appearance* and social presence of VHs, particularly with variable quantities of light from the physical background and from the VH in the foreground?

*e-mail: juanita.benjamin@ucf.edu
†e-mail: bruder@ucf.edu
‡e-mail: carsten.neumann@ucf.edu
§e-mail: dirk.reiners@ucf.edu
¶e-mail: carolina@ucf.edu
‖e-mail: welch@ucf.edu

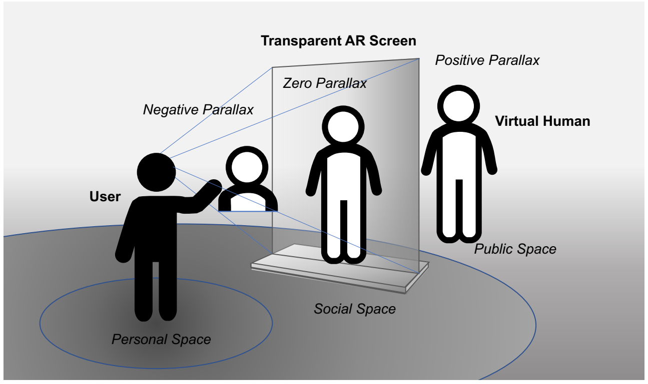


Figure 1: Illustration of a virtual human installation with a transparent AR screen, which acts as a window through which users can see and interact with a virtual human, similar to the vision introduced in the movie *The Time Machine (2002)*. The virtual human can be displayed stereoscopically in front of (*negative parallax*), centered at (*zero parallax*), or behind (*positive parallax*) the screen. Note that transparent screens use an additive light model, which means that the appearance of the virtual human depends on the amount of light in the *foreground* (virtual human on the screen) and in the *background* (physical environment behind the screen).

RQ2 How does the distance between the AR display surface and the participants, as well as the quantity of foreground and background light, affect *proxemics* with VHs?

2 Background

In this section, we provide background information on the technologies and research questions addressed in this paper.

2.1 AR Transparent Displays

Transparent screens are an emerging technology that can overlay two- or three-dimensional virtual content over the view of a physical environment [29, 37, 39]. These screens can be realized using a variety of technologies, including front-projection or rear-projection screens, transparent LCDs, etc. [4, 22, 42]. Furthermore, three-dimensional content may be conveyed using polarization or shutter glasses, or auto-stereoscopic or light-field displays [7, 8, 38]. In our work, we used an isotropic rear-projection screen and shutter glasses. Isotropic screens diffuse all incoming light from any direction as opposed to anisotropic screens, which redirect light coming from a specific direction [39].

Transparent screens present imagery based on an *additive light model*. Using this model, light coming from the display in the foreground is blended with light from its physical background [11,33, 34]. As a result, virtual imagery appears transparent if the contrast between the foreground light (coming from the display) and the background light (from the physical environment) is too low [12, 16, 27, 32, 44]. Poor contrast makes it difficult for an observer to accurately perceive the exhibited virtual imagery [10, 31, 45]. In that sense, transparent screens are very similar to OST HMDs though they are different in other aspects. For instance, with OST HMDs, a user's accommodation of their eyes is usually fixed to its focal distance, whereas with transparent screens, a user's accommodation depends on their distance from the physical screen.

Applications where transparent AR screens have been used in the past include telepresence and collaboration [28]. Pluss et al. [38] created a life-sized 3D projection screen for the purpose of telecommunication. It was designed to integrate a remote conversation partner into the local environment of another person, enabling hand gestures and body language. Mock et al. [35] integrated a stereoscopic semi-transparent display with a calibrated stereo camera into a video conferencing application for 3D remote collaboration.

2.2 AR Virtual Humans

In the scope of this work, VHs may either refer to computer-controlled agents or user-controlled avatars. Related work suggested that a VH's appearance may be affected in a similar way by the additive light model of OST displays as other virtual content, i.e., showing a preference for lighter-colored VHs, as darker colors appear more transparent. For instance, Doroodchi et al. [9] evaluated how participants perceive the appearance of their own self-avatar on a Microsoft HoloLens 2 OST HMD. In their study, participants were tasked with modeling their own avatar after themselves while looking in the mirror, by choosing hairstyles, skin color, and clothes. They investigated how participants would represent themselves with the self-avatars under two lighting/transparency conditions (200 and 2,000 lux). Their findings revealed that a lighter physical environment had an effect on how participants chose their self-avatar skin colors, primarily giving avatars (particularly those with darker skin tones) a lighter shade.

Peck et al. [36] further ran a large online study with simulated imagery that focused on the dehumanization of avatars and human faces with different levels of transparency. The participants were presented a series of images showing VHs such as a stereotypically White man, Black man, Asian man, and a zombie. In these images, different levels of transparency of the VHs were present, ranging from 0% to 65%. The participants were instructed to evaluate each image and state whether the VH seemed more human, animal-like, robotic, competent, etc. Additionally, the researchers assessed their perceived emotions, motivated by related work, which found that Black Americans were perceived as angrier than White Americans even when the facial expressions were constant [17, 24]. The participants in this work had to rate how happy or angry each VH appeared. In their results, it was evident that the dehumanization and misperception of emotions occurred when the transparency increased.

2.3 Proxemics in AR

Proxemics denotes humans' use of space and the distances that individuals maintain between each other in social situations. Edward T. Hall defined four zones that include *intimate space*, *personal space*, *social space*, and *public space* [18–20]. Intimate space is reserved for intimate partners, and ranges from 0 m to 0.45 m. Personal space (0.45 m to 1.22 m) is reserved for trusted individuals like friends or family members. Social space (1.22 m to 3.06 m) is the space in which an individual would interact with acquaintances. Lastly, public space (beyond 3.06 m) is mainly used for public speaking.

Many researchers have studied proxemic behaviors in AR/VR environments, and their findings have shown that the appearance or the behavior of the VH impacts the user's distance from the VH [1]. Huang et al. [23] conducted a study in AR where the participants were asked to interact with six VHs (two males, two females, a robot and a pillar) in an art gallery. They found that when approaching these agents and asking for directions, the participants respected the agents' personal space and modulated interpersonal distances according to the human-like agents' perceived gender. Bailenson et al. [2] conducted a study in VR to test interpersonal space and mutual gaze with human agents and non-human objects. Participants were asked to walk around a virtual room, read the names of the agents, and remember them. They found that participants maintained more space around the human agents compared to non-human objects.

Human responses to virtual and real humans vary depending on the situation. If a virtual human behaves awkwardly or lacks interactivity, individuals may feel uncomfortable being too close. Lee et al. [30] studied human locomotion behavior and proxemics in the presence of real or virtual humans in AR. They focused on the limitations of current OST displays and their impact on locomotion behavior near virtual humans. The study found that participants maintained a greater distance, walked a longer path at a slower pace around virtual humans compared to real humans.

3 USER STUDY

In this section we describe the user study we conducted to understand the effects of the AR display technology on the two research questions stated in Section 1.

3.1 Participants

We recruited 22 participants for our user study; 9 female and 13 male (ages 18 to 56, M=28.5, SD=8.0). All participants were members of the local university community. The participants had normal or corrected-to-normal vision; 7 participants wore glasses during the experiment, and 2 wore contact lenses. Apart from one participant, who indicated strong astigmatism (which we did not consider a sufficient reason for exclusion), none of the participants reported any known visual disorders, such as color or night blindness, stereo blindness, or dyschromatopsia. None of the participants reported any known cognitive or motor disorders. With respect to VR, ratings of experience among the participants indicated 1 with limited experience, 6 with some experience, and 15 with expert levels of experience. For AR, ratings of experience indicated 5 participants with limited experience, 6 with some experience, and 11 with expert levels of experience. Three of our participants indicated that they had watched the Science Fiction movie *The Time Machine (2002)*, where a similar setup as described in this paper was shown.

3.2 Material

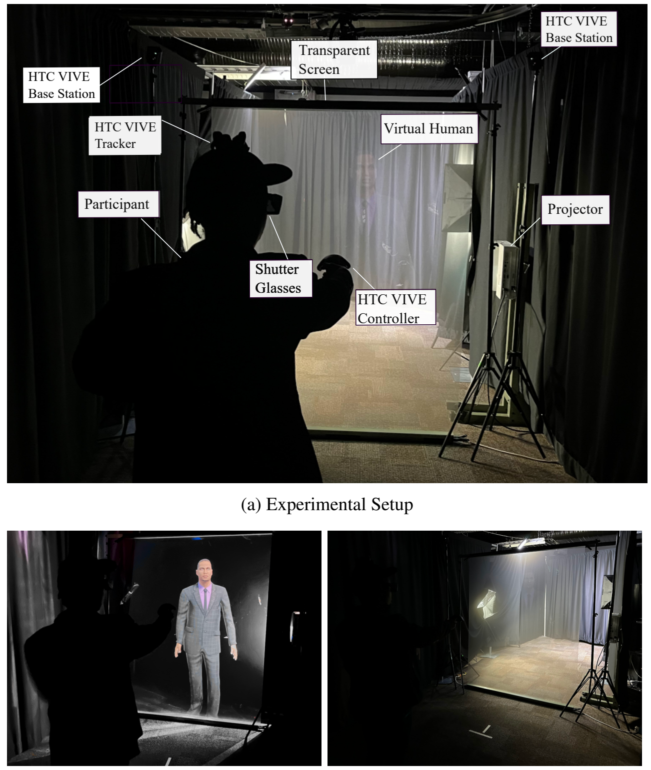
The setup and material we used for our study is described below.

Physical hallway setup. The hallway where participants stood and the transparent screen was mounted was 8 m long and 2.5 m wide. The hallway was framed by opaque black curtains with a height of 2.16 m. The curtains were tall and thick enough to block outside light from reaching the participants. Behind the transparent screen, we mounted four professional (diffusing) photography lights that could illuminate the background behind the screen in two conditions: In the darker condition (see Figure 2b), the amount of light was $<$1 lux throughout the physical background. In the lighter condition (see Figure 2c), the amount of light in the physical background ranged from 235.7 lux on the carpeted floor to 198.6 lux on the black curtains in the back. We measured the amount of light with an Urceri MT-912 Light Meter. This light meter is reported to have an accuracy of $\pm$3% of the measured value, and can make measurements between 1 and 200,000 lux, as reported on the Urceri website[1]. The amount of light we chose for our lighter physical background fell in the range of typical office lighting (200–500 lux)[2].

Transparent screen. The screen consisted of a transparent foil called Sunice Rear Projection Film[3], which provides high color reappearance due to its high light transmittance and diffusion. It has a thickness of 100 μ, transmittance percentage of 54%, a haze of 33.5%, and a viewing angle of 120 degrees. The screen surface was 1.5 m wide and 2 m tall. The top of the foil was attached to a fiber glass mount. There were small holes on each side of this fiber glass mount where two aluminum stands were attached. Another fiber glass mount was attached to the bottom to keep it stretched.

Projector. The projector used in this experiment was an Optoma EH330UST with a resolution of 1920×1080, Full HD 3D, providing 20,000:1 contrast at 3,600 ANSI lumen. The projector was placed 0.6 m behind the transparent screen and 0.6 m to the right. To view stereoscopic content, the projector was set to 3D mode with DLP Link, a Volfoni ActivHub Radio Frequency (RF) emitter was used, and the participants wore Volfoni EDGE RF shutter glasses. The projection was calibrated in Unity using our in-house calibration tools.

[1] https://www.urceri.com/mt-912-light-meter.html
[2] https://www.ppsthane.com/blog/lux-level-standards-industry
[3] https://www.aliexpress.us/item/3256801637858741.html



(a) Experimental Setup

(b) Darker Background ($<$1 lux) (c) Lighter Background (198.6–235.7 lux)

Figure 2: Experimental setup: (a) Annotated photo showing a participant in the experiment setup with a VH presented on a transparent projection screen mounted in a hallway (framed by curtains) within the laboratory space. The participant is wearing shutter glasses and holding a tracked HTC VIVE controller. The participant's head is tracked through an HTC VIVE (Lighthouse V2) tracking sensor attached to a baseball cap. The setup with (b) darker and (c) lighter background light.

Visual stimuli. We chose two types of VHs, with two levels of skin colors, while both were dressed in professional attire. We took these models from the Microsoft RocketBox library[4]. These models were then uploaded to Mixamo[5], where we applied an idle animation. The models had blendshapes for the three facial expressions we used in this experiment: angry, happy, and neutral. We chose these three facial expressions due to related work by Peck et al. [36], who suggested that transparent VHs may appear more angry/aggressive than opaque VHs. For rendering, we used the Unity Engine version 2021.3.19f1 LTS, and ran it on an Alienware 17 R4 computer with Windows 11 Pro, Intel Core i7-8750, 32 GB RAM, and NVIDIA GTX 1070 GPU.

Tracking and controllers. For tracking purposes, we used two HTC VIVE Base Stations (Lighthouse V2) that were positioned on either side of the transparent AR screen. For head-perspective rendering, participants wore a baseball cap with an HTC VIVE tracker attached. In addition, each participant used their dominant hand to hold a tracked HTC VIVE controller.

3.3 Methods

To answer our two research questions (see Section 1), we split our user study into two experiments. The same participants took part in both experiments, which were conducted one after another with a short break.

[4] https://github.com/microsoft/Microsoft-Rocketbox
[5] https://www.mixamo.com

(a) Lighter physical background (198.6–235.7 lux); lighter VH skin color (7.7 lux) (b) Lighter physical background (198.6–235.7 lux); darker VH skin color (5.9 lux) (c) Darker physical background (<1 lux); lighter VH skin color (7.7 lux) (d) Darker physical background (<1 lux); darker VH skin color (5.9 lux)

Figure 3: Photos showing the different lighting conditions: (a-b) lighter and (c-d) darker physical background; VH with (a, c) lighter and (b, d) darker human skin color. Note the differences in transparency caused by the different amounts of background and foreground light.

(a) Darker VH skin color; happy expression (b) Darker VH skin color; neutral expression (c) Darker VH skin color; angry expression

(d) Lighter VH skin color; happy expression (e) Lighter VH skin color; neutral expression (f) Lighter VH skin color; angry expression

Figure 4: Photos of the VHs with different amounts of foreground light and facial expressions: (a-c) darker and (d-f) lighter VH skin color; (a,d) happy, (b,e) neutral, and (c,f) angry facial expression.

3.3.1 Experiment I: Appearance

Design. For this part of the study, we used a 2 (background light) × 2 (foreground light) × 3 (VH facial expressions) within-subjects design.

Factors. We investigated the following factors (see Figure 4):

- **Background light (2):** darker (<1 lux throughout) and lighter (max 235.7 lux, min 198.6 lux) physical background behind the transparent AR screen;
- **Foreground light (2):** darker (5.9 lux) and lighter (7.7 lux) foreground VH skin color;
- **VH facial expression (3):** happy, neutral, and angry.

Measures. During these trials, participants were standing 2 meters in front of the transparent AR screen, while the VH was presented centered at the location of the screen. A questionnaire was presented to participants on a laptop that we placed on a podium in front of them. We collected the following subjective responses from participants to measure their perception of the different VH appearances. The semantic differentials were assessed for all three factors *Background Light*, *Foreground Light*, and *VH Facial Expression*, while social presence was assessed only for *Background Light* and *Foreground Light* after participants had seen all three facial expressions:

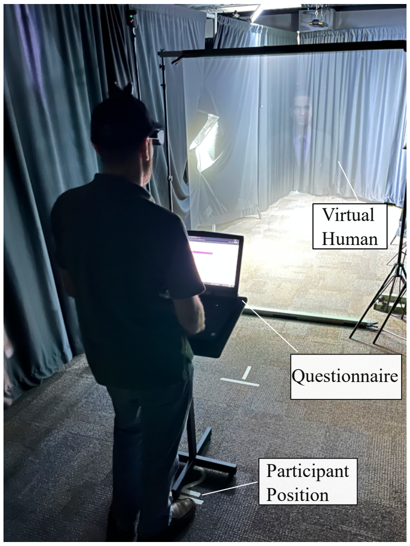
- *Social presence questionnaire*: We asked participants to fill out the standard social presence questionnaire introduced by Bailenson et al. [3]. It consists of five questions, which are scored on 7-point Likert-like scales (from −3 to +3). We computed the overall social presence score as the mean of these five questions, with a positive score indicating a high sense of social presence and a negative score indicating a low sense of social presence [3].
- *Semantic differentials*: Inspired by Peck et al.'s work [36], the following semantic differentials [43] were presented, asking participants to rate the appearance of the VH on a 7-point scale:
  - S_1: 1=sad to 7=happy
  - S_2: 1=angry to 7=calm
  - S_3: 1=machine to 7=human
  - S_4: 1=creepy to 7=pleasant

3.3.2 Experiment II: Proxemics

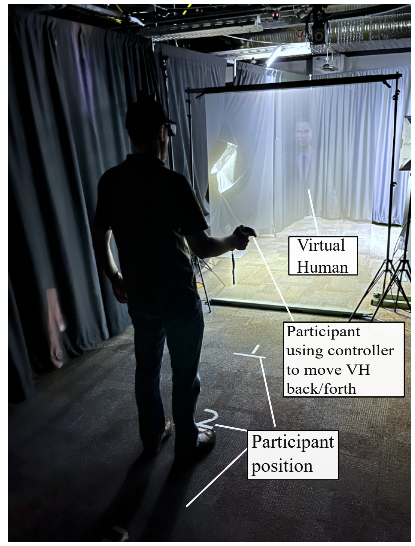
Design. For this part of the study, we used a 2 (background light) × 2 (foreground light) × 3 (screen distances) × 2 (proxemics thresholds) full-factorial within-subjects design.

Factors. We investigated the following factors (see Figure 3):

- **Background light (2):** darker (<1 lux throughout) and lighter (max 235.7 lux, min 198.6 lux) physical background behind the transparent AR screen;
- **Foreground light (2):** darker (5.9 lux) and lighter (7.7 lux) foreground VH skin color;
- **Screen distance (3):** Distances of 1, 2, and 3 meters at which participants stood in front of the transparent AR screen;
- **Proxemics threshold (2):** Participants estimated the following thresholds in the scope of Hall's theory of proxemics [18, 19] (see Figure 1): the threshold from *personal space* to *social space*, and the threshold from *social space* to *public space*; together these two thresholds indicate the start and end of the *social space* in which most interpersonal interactions take place.



(a) Experiment I: Appearance



(b) Experiment II: Proxemics

Figure 5: Annotated photos showing the tasks participants completed in the two experiments during the user study: (a) While standing 2 meters in front of the transparent AR screen, and the VH was presented at the distance of the screen, the participant rated the VH's appearance with a questionnaire on a laptop that was elevated on a podium in front of them. (b) For the proxemics experiment, participants used the up/down buttons on the HTC VIVE controller to move the VH in front of them closer or farther away from them until they estimated the VH to be at the right distance, at which point they pulled the trigger of the controller and the next trial started.

Measures. Participants estimated the two proxemic thresholds using the following methodology:

- *Distance adjustment task:* Participants were holding the tracked HTC VIVE controller with their dominant hand. They were instructed to use the up/down buttons on the controller to move the VH gradually closer to or farther away from them. When they felt confident that the VH was at the right distance matching the indicated proxemics threshold, they pulled the trigger on the controller, and the next trial started. No feedback about their accuracy was given to participants.

3.4 Hypotheses

Based on related work in this field (see Section 2) and our own reasoning, we developed the following hypotheses, which we evaluated in this experiment with our subjective and objective measures:

H1 Estimates of the appearance of the VHs will be degraded by higher amounts of *Background Light* and lower amounts of *Foreground Light* with respect to the **(a)** social presence questionnaire and **(b)** semantic differentials.

Reasoning: VHs become harder to see in the presence of transparency caused by a high amount of background light that washes out the foreground, and/or a low amount of foreground light that is insufficient to be visible over the background. Transparency was suggested in related work as a potential source for misperceived emotions [36] and misperceived self-avatar appearances [9].

H2 The transparent AR screen setup will have an effect on proxemic threshold estimates with the VH: **(a)** The larger the *Screen Distance*, the farther participants will indicate the proxemics thresholds from personal to social space, and from social to public space. **(b)** The more *Background Light* and/or less *Foreground Light*, the shorter participants will indicate the proxemics thresholds.

Reasoning: Proxemic distances depend on a variety of cues related to the appearance of the interlocutor, its perceived distance, and the social situation they are in, among others. Re. (a): We anticipated that in the special situation of the transparent AR screen, participants would not only treat the VH as a social entity, but also consider the screen itself as a social situational entity, which could shift their estimates farther out, the farther away the screen is from participants. Re. (b): Our reasoning behind this hypothesis was that if the VH is perceived as more transparent, it becomes harder to see, such that moving them closer to the observer may compensate for their limited visibility.

4 Results

In this section, we present our analysis and results.

4.1 Subjective Responses

We analyzed the semantic differentials and social presence combined scores with repeated-measures analyses of variance (RM-ANOVAs) and Tukey multiple comparisons with Bonferroni correction at the 5% significance level. We understand that there is a debate about the use of parametric tests for this class of responses; we point interested readers to [5]. We confirmed our results both with parametric and non-parametric tests but prefer to report the parametric test results due to their higher expressiveness. For our analysis, we confirmed the normality with Shapiro-Wilk tests at the 5% level and QQ plots. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly's test indicated that the assumption of sphericity was not supported.

4.1.1 Social Presence

The results for our social presence scores in the appearance part of the experiment are shown in Figure 6. We analyzed the results for effects of the *Background Light* and *Foreground Light* factors on the scores. The statistical test results of the two-way RM-ANOVA are shown in Table 1(top).

We found support for our Hypothesis **H1(a)** that the social presence scores were significantly different between the two levels of *Background Light* we tested in the experiment. The scores were significantly higher in the condition with the darker background,

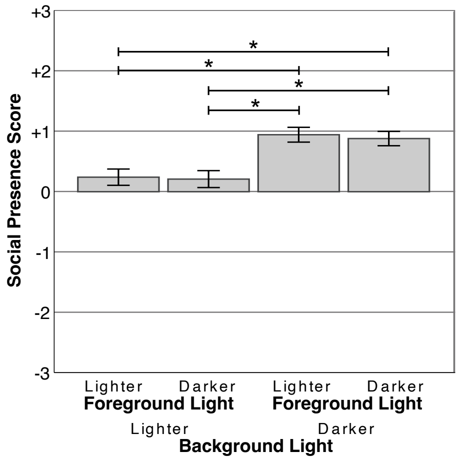


Figure 6: Experiment I Appearance: Results for the social presence scores for the two factors *Background Light* and *Foreground Light*. The *y*-axis indicates the mean scores from Bailenson et al.'s Social Presence Questionnaire from -3 to +3, with a positive score indicating a high sense of social presence and a negative score indicating a low sense [3]. The vertical bars indicate the standard error. The horizontal whiskers and asterisks indicate statistical significance of the pairwise comparisons.

which resulted in less transparency of the VH than the condition with the lighter background. We found no significant difference for our two levels of *Foreground Light*. We discuss these results in Section 5.1.

4.1.2 Semantic Differentials

The results for the four semantic differentials S_1–S_4 in the appearance part of the experiment are shown in Figure 7. We analyzed the results for effects of the *VH Facial Expression*, *Background Light*, and *Foreground Light* factors on the scores. The statistical test results of the three-way RM-ANOVAs and pairwise comparisons are shown in Table 1. Our results passed our sanity check, meaning that we found a significant difference between the three facial expressions and the first two semantic differentials: the VH's happy facial expression was rated as more happy in S_1, and the angry facial expression was rated as more angry in S_2. We also observed that participants rated the appearance of the VHs with a neutral facial expression as more pleasant than both happy/angry facial expressions.

However, we found no support for our Hypothesis **H1(b)** that the semantic differentials for the appearance of the VH were affected by the *Foreground Light* or *Background Light*, with scores that were overall very similar. In other words, even severe transparency did not have a noticeable effect on how participants estimated the VHs' emotions (S_1 and S_2), humanness (S_3), and creepiness (S_4). We discuss these results in Section 5.1.

4.2 Objective Responses

We analyzed the objective responses (proxemics thresholds) with RM-ANOVAs and Tukey multiple comparisons with Bonferroni correction at the 5% significance level. We confirmed the normality with Shapiro-Wilk tests at the 5% level and QQ plots. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly's test indicated that the assumption of sphericity was not supported.

4.2.1 Proxemics Estimates

The results for the distance adjustment task in the proxemics part of the experiment are shown in Figure 8. We analyzed the results for effects of the *Background Light*, *Foreground Light*, *Screen Distance*, and *Proxemics Threshold* factors on the proxemic distances. The statistical test results of the four-way RM-ANOVA and pairwise comparisons are shown in Table 2. Our results passed our sanity check, meaning that we found a significant difference between the two tasks related to the *Proxemics Thresholds* we asked our participants to indicate.

In line with Hypothesis **H2(a)**, our results show a significant main effect for *Screen Distance* on the proxemics distances. As can be seen in Figure 8(a), the farther away participants stood from the transparent AR screen, the farther away they placed the VH for both proxemics thresholds, i.e., the entire *social space* was shifted backwards. For a 1-meter distance from the transparent AR screen, the social space ranged from 1.23 to 3.06 meters (1.83 m long); for a 2-meter distance, it ranged from 1.49 to 3.43 meters (1.85 m long); and for a 3-meter distance, it ranged from 1.75 to 3.86 meters (2.11 m long). We discuss our results in detail in Section 5.2.

We found no support for our Hypothesis **H2(b)** that the proxemics thresholds in our study were affected by the *Foreground* and *Background Light*, showing results that were overall very similar (see Figures 8b and 8c). It is interesting that even in the presence of severe transparency, it did not have a noticeable effect on proxemics. We discuss these results in Section 5.3.

5 Discussion

We discuss our main findings in this section. Overall, our results show that transparent AR screens can be an effective means to enable interactions between virtual and real humans in a shared space, though with a few differences compared to interactions between real humans.

5.1 Social presence with the virtual human was significantly reduced by higher transparency, while ratings of emotions and humanness were not

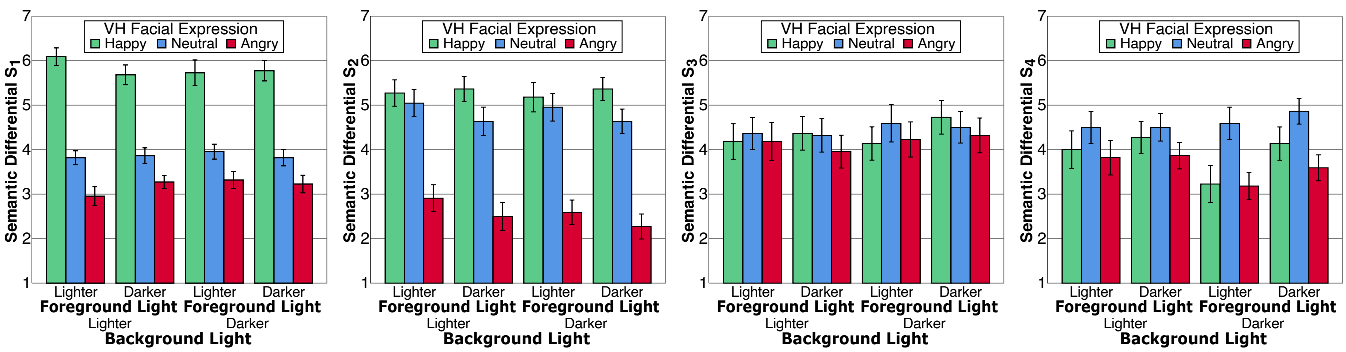
We found support for our Hypothesis **H1(a)** in that our participants indicated a higher sense of social presence with the VHs in the darker *Background Light* condition. We found no effect of *Foreground Light* on the scores, which may be explained by the large difference in transparency caused by the two levels of *Background Light* compared to the minor difference in transparency for the two levels of *Foreground Light*. Our results indicate that a darker environment will be more effective for social interactions between real and virtual humans through a transparent AR screen.

Contrary to our Hypothesis **H1(b)**, we found no evidence in our results for an effect of transparency on our participants' estimates of the VH's emotions (happy or angry), or how human or creepy the VH appears. Neither of our two levels of *Background Light* nor our two levels of *Foreground Light* had a significant effect on how happy/sad, angry/calm, machine/human, or creepy/pleasant the VH appeared, with scores that were overall very similar. When we designed our experiment, we anticipated to find such effects as they were motivated by related work, though using different display technologies. Peck et al. [36] found an effect of transparency on estimated VH appearances based on an online study with simulated 2D (non-stereoscopic) imagery, and Doroodchi et al. [9] found that transparency affected the appearance of a self-avatar shown on an OST HMD. It is interesting that despite some highly transparent conditions, our participants were able to perceive the facial expressions comparatively well. Moreover, there was no noticeable increase in ratings of how non-human or creepy the VHs appeared for the higher-transparency conditions.

Taken together, our results for the semantic differentials and social presence scores indicate no noticeable biases in the perception of the VHs due to the transparent AR display technology but a general sense of lower social presence for high levels of transparency. Future work may confirm our results with a wider range of facial

Table 1: Experiment I Appearance: Statistical test results for the (top) social presence scores and (bottom) four semantic differentials S_1–S_4 (see Section 3.3.1). The latter measure was assessed for all three factors *Background Light*, *Foreground Light*, and *VH Facial Expression*, while the former measure was assessed only for *Background Light* and *Foreground Light* after participants had seen all three facial expressions.

Measure	**RM-ANOVA**	**Factor**	**df**$_G$	**df**$_E$	***F***	***p***	η_p^2	**Pairwise Comparisons**
Social Presence Score	Two-way	Background Light	1	21	16.34	**<0.001**	0.44	N/A
		Foreground Light	1	21	2.14	0.16	0.09	N/A
S_1 Score	Three-way	Background Light	1	21	0.05	0.82	0.002	N/A
		Foreground Light	1	21	0.22	0.64	0.01	N/A
		VH Facial Expression	2	42	138.06	**<0.001**	0.87	All ***p*<0.001**
S_2 Score	Three-way	Background Light	1	21	0.65	0.43	0.03	N/A
		Foreground Light	1	21	1.15	0.30	0.05	N/A
		VH Facial Expression	2	42	75.78	**<0.001**	0.78	All ***p*<0.01** except (happy, neutral)
S_3 Score	Three-way	Background Light	1	21	0.90	0.35	0.04	N/A
		Foreground Light	1	21	0.70	0.41	0.03	N/A
		VH Facial Expression	2	42	2.00	0.15	0.09	All *p*>0.05
S_4 Score	Three-way	Background Light	1	21	1.20	0.29	0.05	N/A
		Foreground Light	1	21	2.90	0.10	0.12	N/A
		VH Facial Expression	1.33	27.93	10.70	**0.001**	0.34	All ***p*<0.05** except (happy, angry)



(a) S_1: 1=sad to 7=happy (b) S_2: 1=angry to 7=calm (c) S_3: 1=machine to 7=human (d) S_4: 1=creepy to 7=pleasant

Figure 7: Experiment I Appearance: Results for the four semantic differentials for the three factors *VH Facial Expression*, *Background Light*, and *Foreground Light*. The *y*-axes indicate the semantic differential score for (a) S_1 (1=sad to 7=happy), (b) S_2 (1=angry to 7=calm), (c) S_3 (1=machine to 7=human), and (d) S_4 (1=creepy to 7=pleasant). The vertical bars indicate the standard error.

expressions and semantic estimates as well as more engaging and longer-term social interactions with the VHs.

5.2 Proxemics thresholds depend on the user's distance from the transparent AR screen

Our proxemics results show that for the closest distance of 1 meter at which participants stood in front of the transparent AR screen, both thresholds from personal to social space (1.23 m), and from social space to public space (3.06 m) are very close to the thresholds reported in the real-world proxemics literature (1.22 m and 3.05 m, respectively [18]), making this the situation that most closely matches what we would expect from regular daily interactions between real humans. It is further a very natural situation that could occur in the real world if we think of the transparent screen as a glass divider between two humans: the participant stood closely in front of the transparent screen, while the VH stood in a close range behind the screen. A similar social situation was also presented in the "library scene" of the movie *The Time Machine (2002)*, where the main actor interacted with a VH standing closely behind a glass screen.

However, we found that the farther away participants stood from the transparent AR screen, the farther away from them they shifted the proxemics thresholds. In effect, they moved the beginning of the social space about 0.26 meters farther away from themselves for each meter they stood away from the screen (starting at 1.23, 1.49, 1.75 m for 1, 2, 3 m screen distances, respectively). There were limited differences in terms of the overall length of the social space (1.83, 1.85, 2.11 m social space length for 1, 2, 3 m screen distances, respectively).

As indicated for Hypothesis **H2(a)** in Section 3.4, we anticipated that participants would not entirely dissociate the stereoscopically displayed VH from the transparent AR screen on which the imagery was displayed. For one, if the VH was presented in front of the screen (at negative parallax), the screen's view frustum from the participant's eye position did not allow them to see the entirety of the VH's body, effectively cutting of larger parts of the lower body of the VH the farther away they stood from the screen (see Figure 1). This may explain why participants shifted VHs in front of the screen farther away, but it cannot explain why they also shifted those VHs farther away that were behind the screen. A potential explanation is that participants considered both the VH and the screen as connected social entities for the purpose of proxemics: if the screen was 1 meter away (i.e., the screen was within their personal space), they indicated proxemics thresholds for the VH similar to the real world, while for the 2 and 3 meter screen distances (i.e., the screen was within their close or far social space), the thresholds for the VH were moved correspondingly farther away. It warrants further research to investigate in how far these transparent AR screens influence perceptions of social situations and interactions with VHs presented on these screens.

Table 2: Experiment II Proxemics: Statistical test results for the distance adjustment task for the two proxemics thresholds.

Measure	RM-ANOVA	Factor	df_G	df_E	F	p	η_p^2	Pairwise Comparisons
Estimated Threshold Distance	Four-way	Background Light	1	21	0.60	0.45	0.03	N/A
		Foreground Light	1	21	1.84	0.19	0.08	N/A
		Screen Distance	1.2	25.2	19.30	**<0.001**	0.48	All ***p*****<0.01**
		Proxemics Threshold	1	21	22.02	**<0.001**	0.51	N/A

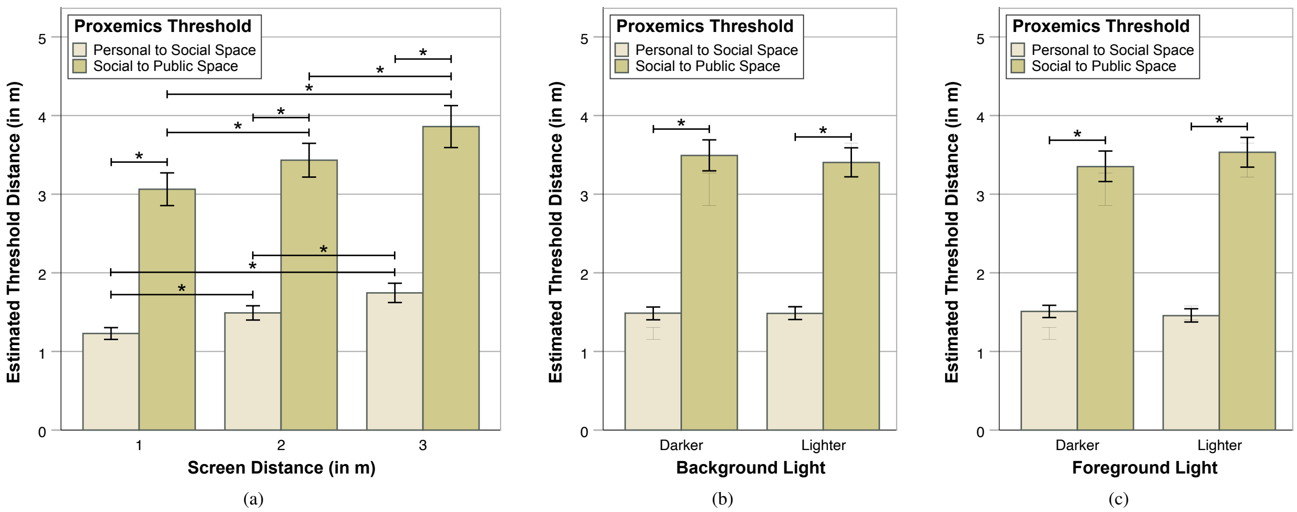


Figure 8: Experiment II Proxemics: Objective responses for both thresholds and the three factors (a) *Screen Distance*, (b) *Background Light*, and (c) *Foreground Light*. The *y*-axes indicate the distance at which participants placed the VH from themselves; the range between the closer threshold to the farther threshold indicates the *social space* range, in which most interpersonal interactions would normally occur. The vertical bars indicate the standard error. The horizontal whiskers and asterisks indicate statistical significance of the pairwise comparisons.

5.3 Proxemics thresholds were not noticeably affected by the transparency of the virtual human

As highlighted in Section 4.2.1, we found no support for Hypothesis **H2(b)** that the transparency of the VH had an effect on proxemics. Neither of the two causes of transparency we considered in this experiment (i.e., *Background* and *Foreground Light*) had a significant effect on the results. When we increased the *Background Light* (max 235.7 lux, min 198.6 lux), the VH became noticeably more transparent (see Figure 3), but participants still estimated the two proxemics thresholds at similar distances as in the conditions where the background behind the screen was darker (<1 lux). Similarly, we found no significant effect when we changed the amount of *Foreground Light* with the resulting slight transparency differences that were caused by choosing a VH with a darker skin color (5.9 lux) compared to a lighter skin color (7.7 lux). Overall, our results suggest that transparency is not a major factor when considering proxemic zones between a real and a virtual human on such transparent AR screens.

5.4 Limitations

We identified three main limitations of our study. First, while we decided to only include male VHs in our study, motivated by similar choices in directly related work by Peck et al. [36], we believe that future work should incorporate a wider and more diverse set of VHs. For instance, a few studies have shown that participants are more inclined to keep a larger distance to male than female VHs, which indicates that gender can have an effect on proxemics [21, 23]. Second, the VHs we chose were not interactive beyond basic idle animations. Higher interactivity might have influenced the participants' level of social presence and could have impacted their judgments of proxemics, which we believe future work should focus on. Third, while we chose the VHs and facial expressions from the Microsoft Rocketbox library, we did not formally verify the accuracy of these facial expressions. None of our participants indicated any issues related to the VHs' facial expressions that we used in our study, but future studies should include a pre-test to verify that participants perceive facial expressions accurately.

6 Conclusion

In this paper, we described a user study in which we evaluated how VHs are perceived when presented on our prototype of a transparent AR screen installation. Our results suggest that the VH's transparency as caused by differences in foreground and background light can have a significant effect on social presence ratings, while the VH's estimated emotions and humanness are comparatively robust to differences in transparency. Our results further showed that proxemics thresholds with the VH differ based on the observer's distance from the transparent AR screen additionally to their distance from the VH. Our results provide important insights for practitioners developing applications based on intelligent VHs and transparent AR displays. Based on our results, future work may look into more complex social interactions and scenarios between real and virtual humans in such shared physical spaces with transparent AR screens.

Acknowledgments

This material includes work supported in part by the National Science Foundation under Award Numbers 2235066 and 1800961 (Dr. Ephraim P. Glinert, IIS); the Office of Naval Research under Award Numbers N00014-21-1-2578 and N00014-21-1-2882 (Dr. Peter Squire, Code 34); and the AdventHealth Endowed Chair in Healthcare Simulation (Prof. Welch).

REFERENCES

[1] F. Argelaguet Sanz, A.-H. Olivier, G. Bruder, J. Pettré, and A. Lécuyer. Virtual proxemics: Locomotion in the presence of obstacles in large immersive projection environments. In *Proceedings of IEEE Virtual Reality (VR)*, pp. 75–80, 2015. 2

[2] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments*, 10(6):583–598, 2001. 2

[3] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis. Interpersonal Distances in Virtual Environments. *Personality and Social Psychology Bulletin*, 29(7):819–833, 2003. 4, 6

[4] S. A. Benton. Holographic displays-a review. *Optical Engineering*, 14(5):402–407, 1975. 2

[5] M. C and A. J. A Comparison of Parametric and Non-Parametric Methods Applied to a Likert Scale. *Pharmacy (Basel)*, 5(2):1–12, 2017. 5

[6] F. Camara and C. Fox. Extending quantitative proxemics and trust to HRI. In *Proc. of the IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 421–427, 2022. 1

[7] B. Delaney. Forget the funny glasses [autostereoscopic display systems]. *IEEE Computer Graphics and Applications*, 25(3):14–19, 2005. 2

[8] N. A. Dodgson. Autostereoscopic 3d displays. *Computer*, 38(8):31–36, 2005. 2

[9] M. Doroodchi, P. Ramos, A. Erickson, H. Furuya, J. Benjamin, G. Bruder, and G. F. Welch. Effects of Optical See-Through Displays on Self-Avatar Appearance in Augmented Reality. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 1–5, 2022. 2, 5, 6

[10] A. Erickson, G. Bruder, and G. F. Welch. Adapting michelson contrast for use with optical see-through displays. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 409–410, 2022. 2

[11] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. A Review of Visual Perception Research in Optical See-Through Augmented Reality. *Proceedings of the International Conference on Artificial Reality and Telexistence & Eurographics Symposium on Virtual Environments (ICAT-EGVE)*, pp. 27–35, 2020. 1, 2

[12] A. Erickson, K. Kim, G. Bruder, and G. F. Welch. Exploring the limitations of environment lighting on optical see-through head-mounted displays. In *Proceedings of the ACM Symposium on Spatial User Interaction (SUI)*, pp. 1–8, 2020. 2

[13] A. Erickson, K. Kim, A. Lambert, G. Bruder, M. P. Browne, and G. Welch. An Extended Analysis on the Benefits of Dark Mode User Interfaces in Optical See-Through Head-Mounted Displays. *ACM Transactions on Applied Perception (TAP)*, 2021. 1

[14] J. Gabbard, M. Smith, C. Merenda, G. Burnett, and D. R. Large. A perceptual color-matching method for examining color blending in augmented reality head-up display graphics. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2020. 1

[15] J. Gabbard, J. Swan, J. Zedlitz, and W. W. Winchester. More than meets the eye: An engineering study to empirically examine the blending of real and virtual color spaces. In *Proceeding of IEEE Virtual Reality (VR)*, pp. 79–86, 2010. 1

[16] J. Gabbard, J. E. Swan, and D. Hix. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence*, 15(1):16–32, 2006. 2

[17] A. G. Halberstadt, V. L. Castro, Q. Chu, F. T. Lozada, and C. M. Sims. Preservice teachers' racialized emotion recognition, anger bias, and hostility attributions. *Contemporary Educational Psychology*, 54:125–138, 2018. 2

[18] E. Hall. *The Hidden Dimension*. Garden City, N.Y.: Doubleday, 1966. 2, 4, 7

[19] E. T. Hall. A System for the Notation of Proxemic Behavior. *American Anthropologist*, 65(5):1003–1026, 1963. 2, 4

[20] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold Jr, M. Durbin, M. S. Edmonson, J. Fischer, D. Hymes, S. T. Kimball, et al. Proxemics [and comments and replies]. *Current Anthropology*, 9(2/3):83–108, 1968. 2

[21] H. Hecht, R. Welsch, J. Viehoff, and M. R. Longo. The shape of personal space. *Acta Psychologica*, 193:113–122, 2019. 8

[22] M. K. Hedili, M. O. Freeman, and H. Urey. Transmission characteristics of a bidirectional transparent screen based on reflective microlenses. *Optics Express*, 21(21):24636–24646, 2013. 2

[23] A. Huang, P. Knierim, F. Chiossi, L. L. Chuang, and R. Welsch. Proxemics for human-agent interaction in augmented reality. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2022. 2, 8

[24] K. Hugenberg and G. V. Bodenhausen. Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, 14(6):640–643, 2003. 2

[25] K. Kim, M. Billinghurst, G. Bruder, H. Duh, and G. F. Welch. Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017). *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 24(11):2947–2962, 2018. 1

[26] K. Kim, A. Erickson, A. Lambert, G. Bruder, and G. F. Welch. Effects of Dark Mode on Visual Fatigue and Acuity in Optical See-Through Head-Mounted Displays. In *Proceedings of the ACM Symposium on Spatial User Interaction (SUI)*, pp. 1–9. ACM, 2019. 1

[27] K. Kiyokawa, M. Billinghurst, B. Campbell, and E. Woods. An occlusion capable optical see-through head mount display for supporting co-located collaboration. In *IEEE International Symposium on Mixed and Augmented Reality*, pp. 133–141, 2003. 2

[28] C. Kuster, N. Ranieri, H. Zimmer, J.-C. Bazin, C. Sun, T. Popa, M. Gross, et al. Towards next generation 3d teleconferencing systems. In *2012 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4. IEEE, 2012. 1, 2

[29] B. Lee and J. Hong. Transparent 3d display for augmented reality. In *Holography, Diffractive Optics, and Applications V*, vol. 8556, p. 855602. SPIE, 2012. 2

[30] M. Lee, G. Bruder, T. Höllerer, and G. Welch. Effects of unaugmented periphery and vibrotactile feedback on proxemics with virtual humans in ar. *IEEE transactions on visualization and computer graphics*, 24(4):1525–1534, 2018. 2

[31] G. E. Legge, G. S. Rubin, and A. Luebker. Psychophysics of reading—v. the role of contrast in normal vision. *Vision Research*, 27(7):1165–1177, 1987. 2

[32] D. Lindlbauer, T. Aoki, R. Walter, Y. Uema, A. Höchtl, M. Haller, M. Inami, and J. Müller. Tracs: transparency-control for see-through displays. In *Proceedings of the annual ACM symposium on User interface software and technology*, pp. 657–661, 2014. 2

[33] M. A. Livingston, J. H. Barrow, and C. M. Sibley. Quantification of contrast sensitivity and color perception using head-worn augmented reality displays. In *IEEE Virtual Reality Conference*, pp. 115–122. IEEE, 2009. 2

[34] M. A. Livingston, J. L. Gabbard, J. E. Swan, C. M. Sibley, and J. H. Barrow. Basic perception in head-worn augmented reality displays. *Human factors in Augmented Reality Environments*, pp. 35–65, 2013. 2

[35] P. Mock, A. Schilling, W. Strasser, and J. Edelmann. Direct 3d-collaboration with face2face-implementation details and application concepts. In *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4. IEEE, 2012. 1, 2

[36] T. C. Peck, J. J. Good, A. Erickson, I. Bynum, and G. Bruder. Effects of Transparency on Perceived Humanness: Implications for Rendering Skin Tones Using Optical See-Through Displays. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 28(5):2179–2189, 2022. 1, 2, 3, 4, 5, 6, 8

[37] S. Peterson. Collimation in transparent projection screens for panoramic augmented reality environments. In *EUROCONTROL Innovative Research and Exhibition Workshop*. Citeseer, 2005. 2

[38] C. Plüss, N. Ranieri, J.-C. Bazin, T. Martin, P.-Y. Laffont, T. Popa, and M. Gross. An immersive bidirectional system for life-size 3d communication. In *Proceedings of the International Conference on Computer Animation and Social Agents*, pp. 89–96, 2016. 1, 2

[39] N. Ranieri, H. Seifert, and M. Gross. Transparent stereoscopic display and application. In *Stereoscopic Displays and Applications XXV*, vol. 9011, pp. 219–225. SPIE, 2014. 2

[40] M. Rehm, E. André, and M. Nischt. Let's come together—social navigation behaviors of virtual and real humans. In *Intelligent Technologies for Interactive Entertainment: First International Conference, INTETAIN 2005, Madonna di Campiglio, Italy, November 30–December 2, 2005. Proceedings 1*, pp. 124–133. Springer, 2005. 1

[41] G. Welch, G. Bruder, P. Squire, and R. Schubert. Anticipating Widespread Augmented Reality: Insights from the 2018 AR Visioning Workshop. Technical report, University of Central Florida and Office of Naval Research, August 6 2019. 1

[42] A. D. Wilson. Touchlight: an imaging touch screen and display for gesture-based interaction. In *Proceedings of the 6th International Conference on Multimodal Interfaces*, pp. 69–76, 2004. 2

[43] Y. M. Yusoff, I. Ruthven, and M. Landoni. The fun semantic differential scales. In *Proceedings of the International Conference on Interaction Design and Children*, pp. 221–224, 2011. 4

[44] L. Zhang and M. J. Murdoch. Perceived transparency in optical see-through augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 115–120, 2021. 2

[45] S. Zuffi, C. Brambilla, G. Beretta, and P. Scala. Human computer interaction: Legibility and contrast. In *Proceedings of the IEEE International Conference on Image Analysis and Processing (ICIAP)*, pp. 241–246, 2007. 2