

Comparing Flat and Spherical Displays in a Trust Scenario in Avatar-Mediated Interaction

Ye Pan, William Steptoe, Anthony Steed

Department of Computer Science, University College London, UK
y.pan, w.steptoe, a.steed@cs.ucl.ac.uk

ABSTRACT

We report on two experiments that investigate the influence of display type and viewing angle on how people place their trust during avatar-mediated interaction. By monitoring advice seeking behavior, our first experiment demonstrates that if participants observe an avatar at an oblique viewing angle on a flat display, they are less able to discriminate between expert and non-expert advice than if they observe the avatar face-on. We then introduce a novel spherical display and a ray-traced rendering technique that can display an avatar that can be seen correctly from any viewing direction. We expect that a spherical display has advantages over a flat display because it better supports non-verbal cues, particularly gaze direction, since it presents a clear and undistorted viewing aspect at all angles. Our second experiment compares the spherical display to a flat display. Whilst participants can discriminate expert advice regardless of display, a negative bias towards the flat screen emerges at oblique viewing angles. This result emphasizes the ability of the spherical display to be viewed qualitatively similarly from all angles. Together the experiments demonstrate how trust can be altered depending on how one views the avatar.

Author Keywords

Spherical displays; trust; telecommunication; avatars; mixed reality

ACM Classification Keywords

H.4.3 Communications Applications: Computer conferencing, teleconferencing, and videoconferencing; H.5.1 Information Interfaces and Presentation: Multimedia Information Systems—*Artificial, augmented, and virtual realities*

General Terms

Design, Experimentation, Human Factors.

INTRODUCTION

In many conversations, such as advice giving or financial transactions, trust is required in order for people to work together effectively [5]. Developing trust relies on a number of

nonverbal cues, of which eye contact seems to stand out [1, 19]. However, traditional video conferencing systems rely on a camera displaced relative to the image of the remote participant, which leads to an immediate misalignment and loss of eye contact [6]. Thus, it is difficult for video conferencing systems to preserve correct gaze direction to enable natural lines of sight operating within Chen's offset threshold [8]. It is even more difficult to allow distant parties to take various seating positions while reproducing correct gaze direction during conversation [35]. A variety of video conferencing systems have been developed to support gaze awareness and improve the level of trust for participants observing from multiple viewpoints (see e.g. [18, 20, 28, 32]). However, those systems usually use flat displays which are visible from only the front.

In this work, we captured a remote person's interpersonal cues and represented them as an animated avatar head on a spherical display. We developed a novel view dependent rendering method for a non-planar display. We integrated with Faceshift[®] enabling a remote person to drive facial expressions of the avatar in real time. We used a spherical display, as it is technically quite simple to build and can be constructed very cheaply in comparison to volumetric displays [13], robotics [10] and animatronic shader lamp avatars [16]. This display is small enough to situate almost anywhere in a room, and it is visible from all directions.

Many researchers have investigated how to maintain or increase levels of trust [19]. However, it is also crucial to ensure that users are able to place trust correctly, avoiding overestimation of the trustworthiness of others [26]. As reviewed in the next section, we have followed the previous work [26] that has conceptualised trust in terms of individuals' choice behaviour in a user-adviser relationship. We investigate two predictions regarding the effect of display type and viewing angle on trust: the spherical display may result in positive bias (i.e. more trust) because it increases social presence; or it may result in better discrimination between trustworthy and less trustworthy actors as it conveys more information. In our two experiments, participants were required to attempt to answer thirty difficult general-knowledge questions over three different viewing angles. For each question, participants could ask for advice from one of two advisers. Unknown to the participants, one was an *expert* who responded with mainly correct information, and the other was a *non-expert* who provided mainly incorrect information. We measured participants' advice seeking behavior as an indicator of their trust in the ad-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI'14, April 26–May 1, 2014, Toronto, Canada.
Copyright © 2014 ACM 978-1-4503-2473-1/14/04...\$15.00.
<http://dx.doi.org/10.1145/2556288.2557276>

viser. The first experiment explores how interpersonal cues of expertise presented on two identical flat displays with different viewing angle affect trust. Results demonstrate that participants were able to discriminate correct advice, but their sensitivity to correct advice decreased at off-center viewing angles. The second experiment compares two display types by investigating how people place their trust. Balanced over participants, the expert appeared either on the sphere or on the flat display. We found most participants preferred seeking advice from the expert, but we also found a tendency for seeking advice from the adviser on the spherical display instead of flat display, in particular when viewed from off-center directions.

In the following sections, we review related work and present the implementation of our spherical display teleconferencing system. This is followed by the evaluation design and two experiments. Finally, we present discussions of the results, conclusions and future work.

RELATED WORKS

Trust in teleconferencing

Trust can be defined as a ‘willingness to be vulnerable, based on positive expectations about the actions of others’ [17]. It plays an important role in interpersonal communication, sometimes even as an enabler for effective communication.

Previous research indicates that it is hard to build trust in teleconferencing, because non-verbal cues were unavailable to be ‘read’ [31]. To determine the effect of eye contact in video-mediated communication on trust, Bekkering and Shim [4] created a scenario in which participants indicated the trustworthiness of a message delivered by people. Results revealed that videos that did not support eye contact resulted in lower perceived trust scores, compared to videos that enabled eye contact. Voice-mail enabled just as much trust as the video that created eye contact, perhaps because lack of eye contact cannot be perceived in audio-only communication. Nguyen and Canny [19] proposed a multiView video-conferencing system and demonstrated that a video-conferencing system that affords more eye contact than the traditional video-conferencing system will create group trust levels similar to those seen in face-to-face group meetings.

As a measure of trust, a popular experimental paradigm currently employed by researchers has been social dilemma games based on the Prisoner’s Dilemma, such as the Day-trader game [9]. Social dilemma games vary in how difficult they are depending on the exact rules and payoff structure, but it generally takes some amount of time and some communication in order to reach the required level of trust [7, 19, 24]. Those games are good models for synchronous and symmetric trust situations, such as two-way conversations. However, in some everyday trust situations, we can identify a trustor who decides first and a trustee who then decides to fulfil or defect, such as one-way conversation. Trust games are suitable models of such situations [25]. Many researchers have investigated the relationship between trust and advice seeking behavior [3, 30]. Riegelsberger et al. [26] investigated user’s trust in advisers and effects of media bias in different representations by observing participants’ advice seeking behav-

ior. In their scenario, participants were asked to participate in a quiz and financial incentives were given for good performance. The questions included in the quiz were extremely difficult, so that good performance required seeking advice. Participants had two advisers but could only ask one for each question. Thus asking one adviser rather than the other can be understood as an indicator of trusting behavior. They found users’ preference for receiving video advice led them to disregard better text-only advice.

Most of this previous work is focused on 2D planar displays. Trust formation on non-planar displays has not been evaluated yet. In pursuing this question, we adapt previous studies of trust to evaluate the advantage of a sphere display over a flat display.

Situated displays

The 360° visibility of situated displays are interesting for teleconferencing, as observers are potentially able to understand the content from any perspective. Many situated displays have achieved accurate reproduction of nonverbal cues, including gaze and deictic gestures.

Animatronic Shader Lamps Avatars (SLA) [16] use the technique where an image is projected onto a screen whose shape physically matches the image itself. It uses cameras and projectors to capture and map the dynamic motion and appearance of a real person onto a humanoid animatronic model. It successfully delivers a dynamic, real-time representation of the user to multiple viewers. The SLA’s use of front projection to texture the 3D facial geometry makes it less practical than one with internal (rear) projection, as commented by the authors. A spherical display offers flexibility compared to humanoid robotics as it isn’t constrained to a single head size or shape. The SphereAvatar [21] system demonstrated a view dependent way of rendering a remote person’s head as an avatar. By tracking the viewer’s position, the display allows the viewer to detect the avatar’s head gaze direction from any viewpoint. Similarly, a previous work used an array of cameras to capture a remote user, and then display video of that person on a spherical display, allowing observers to accurately judge where the remote user was placing their gaze [22]. TeleHuman [15] was a cylindrical 3D display portal for life-size human telepresence. It supported 360° motion parallax as the viewer moves around the cylinder and optionally, stereoscopic 3D display of the remote person. The system successfully preserved gaze and hand pointing cues. It also conveyed 3D body postural information.

Whilst the 360° and gaze preserving capabilities of these situated displays have been demonstrated by these previous works, it is not clear whether there is an effect on the observer’s trust formation from different viewing directions. Our study looks at the effects of spherical display on trust formation at three different viewing angles.

Teleconferencing, avatar-mediated communication

Video conferencing is able to represent participants’ appearance across a distance. However, even minor physical movement of a user may introduce parallax between camera position and video display resulting in loss of gaze awareness.

The 2D nature of standard video constrains the rich spatial cues common to collocated interaction such as depth, resolution, and field of view [29]. In the regard to spatiality, videoconferencing has proven to be more similar to audio conferencing than to unmediated interaction [34].

Recently, avatar-mediated communication, where a remote person is represented by a graphical humanoid, has increased in prevalence and popularity as an emerging form of visual remote interaction [12]. The avatar represents the presence and activities of a remote user and can be visualized using standard displays or projection surfaces in the local room with perspective-correct graphical rendering via head tracking of the local user [27]. Avatars are capable of eliciting appropriate responses from observers (see e.g. [2], [29]). Faceshift[®] demonstrated a high-fidelity and real-time parametric reconstruction of facial expression method using a single depth camera [33]. We have decided to represent a remote user as an avatar instead of video in our experiment, as 3D models are simple to render from any viewing angles.

SYSTEM DESIGN

In the *remote room*, the facial expression of the remote person, the *actor*, is captured. In the *local room*, a single PC renders an animated avatar on a spherical display which is seen by an observer. Figure 1 depicts the system design. We integrated with Faceshift[®] to allow an actor to control the facial expressions of the avatar. We developed a view dependent graphical representation to fully support rendering spherical display surfaces.

Real time facial expression tracking with Faceshift

In the remote room, the actor is recorded in a natural environment using a non-intrusive, commercially available Microsoft Kinect[™]. The actor was seated at the same height of the sensor, about 600 mm horizontal distance from the sensor (see Figure 1). The Microsoft Kinect[™] captures a 640×480 2D color image and a 3D depth map at 30 Hertz, based on invisible infrared projection. It provides a simple and low cost way for acquisition, without the use of face markers, intrusive lighting, or complex scanning hardware.

We used Faceshift[®] with Microsoft Kinect to obtain our actor's facial performances in realtime. Faceshift ensures robust processing given the low resolution and high noise levels of the input data. The output of the tracking optimization is a continuous stream of blendshape weight vectors that drive the avatar. With the embedded plug-in of Faceshift in Maya[®], we obtain 46 blendshapes of the Rocketbox[®] avatar. Then, we export them as .obj format to be used in the ray tracing stage discussed below. Finally, we represent facial expressions as a weighted sum of blendshape meshes, enabling actor to control the facial expressions of the avatar.

View dependent rendering for spherical display

In the local room, the spherical display was the commercially available Magic Planet display by Global Imagination[®]. The Magic Planet is a projection display device with a 16" sphere-shaped surface, a standard 1024×768 projector at 60Hz, and

an internal fisheye lens to project imagery on to the inside of the sphere.

We developed a view dependent rendering method to create 3D object presenting onto spherical image surface, map from the spherical image surface into 2D image plane, and re-project onto spherical display, as if the object is situated inside the sphere display (See Figure 2). The use of a ray tracing engine should provide higher quality images with less distortion than the polygonal rendering approach that was developed for SphereAvatar [21].

We used the NVIDIA[®] OptiX ray tracing engine [23], to trace the path of light from observer's eye to the 3D object through pixels in the spherical image surface.

To implement the ray tracer, we translate the 3D spherical surface into 2D image plane, to represent the surface of the sphere on a flat paper map or on a computer screen. The position of each point (Q) on the 2D image plane (see Figure 2(c)) can be defined by a radius (r , $0 \leq r \leq 1$) and a longitude angle (α , $0 \leq \alpha \leq 2\pi$). The corresponding position of that point (P) on the spherical surface (see Figure 2(a)) can be defined by a latitude angle (θ , $0 \leq \theta \leq \pi$) and longitude angle (ϕ , $0 \leq \phi \leq 2\pi$). The 2D image projector and the display surface are axially symmetric about the optical axis. Thus, the polar angle (α) in 2D image plane is the same as the longitude angle (ϕ) in the 3D spherical surface, shown the Equation 1. All the points at a given radius (r) in the 2D image plane are projected onto the sphere display surface at the same latitude angle (θ). Because of lens distortion, there is a nonlinear relationship between latitude angle (θ) of sphere display surface and the radius (r) of the 2D image plane. We sampled latitude angle (θ) at every 15° to find out the corresponding radius (r) value, shown in Figure 2(b). We use the Matlab[®] second order polyfit to simulate a continuous function as a model to characterizing the relationship between the latitude angle (θ) and the radius (r), presented in Equation 2. Therefore, if we want to project a certain image onto sphere display surface, the corresponding source image can be determined by applying the inverse function to that image.

$$\alpha = \phi \quad (1)$$

$$r = -0.0806 \times \theta^2 - 0.0704 \times \theta + 1.0022 \quad (2)$$

Finally, the 2D image plane that produced above would be projected through the fisheye lens of the sphere display. We could then see a corrected image presented on the spherical surface. In Figure 2(c), the red circle of the 2D image plane corresponds to the equator of the sphere; the center of the 2D image plane projects to a single point on the top of the sphere; the very outer circle in the 2D image plane projects to a single point on the bottom of the sphere. The projected result on sphere display is presented in Figure 2(d).

We use this view dependent graphic representation method discussed above to ray trace an avatar's upper body and head. Figure 3(a) to Figure 3(g) present some sampled mapping results in 2D image plane generated at different viewer's positions while the avatar is looking at the front. Once projected

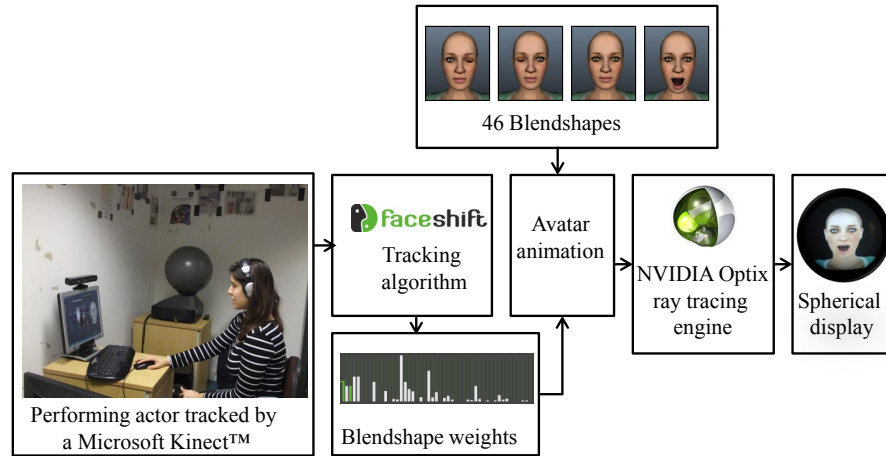


Figure 1. Pipeline for representing an avatar with dynamic facial expressions controlled by an actor on the spherical display.

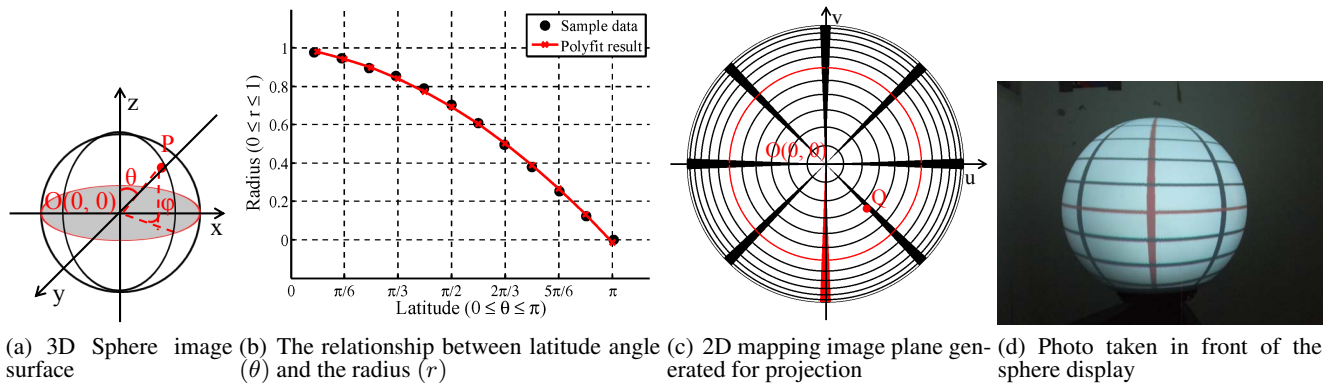


Figure 2. The mapping relationship: each point P on the 3D spherical surface in the subfigure (a) translates into corresponding point Q on the 2D image plane in the subfigure (c), according to calibrated relationship in the subfigure (b). The subfigure (d) shows the projected result of the 2D image plane.

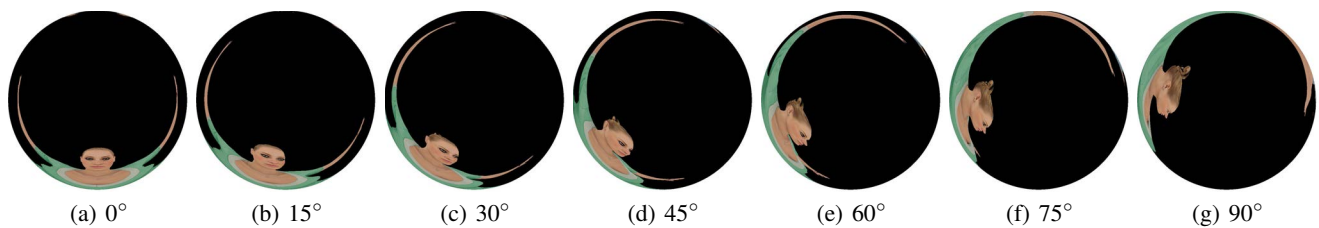


Figure 3. 2D mapping image generated for projection at different viewer positions.



Figure 4. Photo taken at approximately 45° left side of sphere display. For both subfigure (a) and (b), the viewers' positions are the same as the photo taken position. The avatar head is looking at the viewer in the subfigure (a), but the avatar head is looking at the right of the viewer in the subfigure (b). For subfigure (c) and (d), each viewer's position is at right and left side of the photo taken position, respectively.

through the fisheye lens of the display, such images would appear as correctly shaped head and upper body (see Figure 4). This method successfully avoids any seams, overlaps or registration errors in the resulting composite image in projected image on the sphere display. It is also could be extended to other display systems that have a three dimensional display surface.

EVALUATION DESIGN: ADVICE SEEKING BEHAVIOR

Through two experiments, we investigated how display types affect trust. Our first experiment (E1) explored the effect of viewing angle on trust in traditional flat displays, and provided a benchmark by which to measure the spherical display. Our second experiment (E2) investigated the impact of the spherical display given that it could faithfully reproduce the actor's gaze at all viewing directions.

We modeled our experiments on a user-adviser relationship [26], a widely used research paradigm in social psychology. Participants were asked to answer thirty difficult general-knowledge questions and they received chocolates depending on their performance. We gave participants two advisers presented on two teleconferencing displays. Unknown to participants, the two advisers are with different levels of expertise. Additionally, the spatial arrangement of participant-to-displays was varied over the course of the experiment, thereby manipulating participants' viewing angle of the advisers. Advice was free, but only one adviser could be asked per question.

We measured participants' advice seeking behavior under risk as an indicator of trust in the adviser. People generally decide to trust others when facing situations involving risk and uncertainty [11, 17]. Uncertainty arises from the fact that the participants cannot directly observe the two advisers' ability (e.g. expertise) and motivation (e.g. desire to deceive). They need to infer those from interpersonal cues, as the questions were extremely difficult. When recording the non-expert clips, the actor exhibited less direct eye contact and less confident facial expression. When recording the expert clips, the actor exhibited confidence through more positive facial expression, such as smiles and eye contact. In our experiments, viewing angles and display types influence those interpersonal cues. Seeking advice from one adviser in preference over the other could be an indication of trust in that adviser, because receiving poor advice carried the risk of missing out better advice and therefore the participant was less likely to get the correct answer.

Apparatus and materials

Questions

For E1 & E2, we used 30 questions and answers and a transcript of advice from previous research on trust in a human adviser [26]. Those questions are difficult general knowledge questions, to minimize effects of participants' prior knowledge. Examples of questions that were included are 'Which New York Building featured a mural depicting Lenin?' and 'Which one of these is a coastal city in North Korea?'. Based on the pre-test results, the mean probability for pre-testers giving a correct answer was .31 (SD = .11). This value was

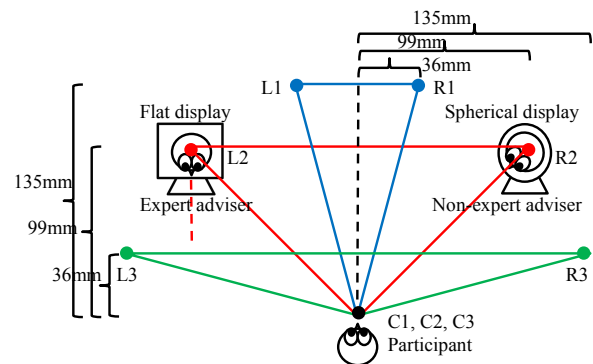


Figure 5. Schematic layout of experiment setup. L1, R1 & C1; L2, R2 & C2 and L3, R3 & C3 are three participant-to-displays spatial arrangements. C1, C2 and C3 are participants' seating positions which are 75°, 45° and 15° relative to display, respectively. Also see Figure 6.

only marginally above chance (.25), indicating that very difficult questions that had been picked.

Expertise

The non-expert and expert advisers were created by recording advice from the same individual before and after training, respectively. The same animations are used in both experiments. We used Faceshift to simultaneously record the actor's performance including voice and blendshape weight vectors that drive the avatar's facial expression. Then, we synchronously replayed both audio and facial expression on the display. The expert and non-expert advisers only differed in the ratio of correct to incorrect advice and in their cues to confidence about the answers. As each time the observer only had access to one of the advisers, they were unaware that both advisers were in fact the same individual recorded at different levels of expertise. For the non-expert adviser, the proportion of correct (i.e. confident) advice was 0.36. For the expert adviser, the proportion of correct (i.e. confident) advice was 0.80. Two incorrect (and less confident) pieces of advice from the untrained recording were added to the expert, in order to avoid artificial perfection.

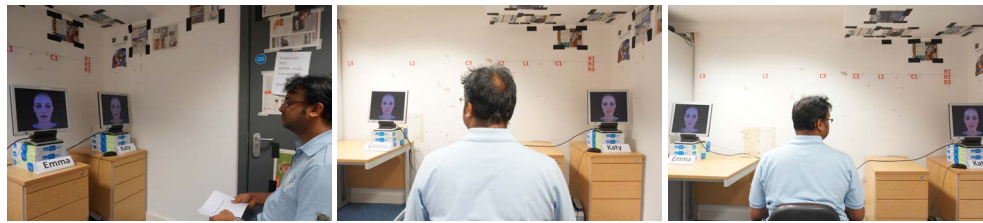
Note that the system as designed and built is a realtime collaborative system that can connect a remote room to a local room. For the purposes of our controlled experiment we used pre-recorded clips.

Display Type

The participant observes the pre-recorded avatar video clips on two displays. We used two flat displays in E1, whereas one flat display and one sphere display in E2. For the flat display, a conventional PC screen was used with a resolution of 1024×768 pixels. For the sphere display, with perspective-correct ray traced imagery, the participant perceives the avatar to be situated inside the display and looking at him or her. We ensured the avatars' apparent sizes on sphere and flat display were the same (20 cm in height).

Seat Position & viewing angle

For both experiments, we arranged the two displays and participant's seat positions at vertices of three isosceles triangles with base angle of 75°, 45° and 15° for three different seat



(a) E1: Photo taken at left side of participant seated at C1 (b) E1: Photo taken behind participant seated at C2 (c) E1: Photo taken slightly right behind participant seated at C3



(d) E2: Photo taken slightly right behind participant seated at C1 (e) E2: Photo taken behind participant seated at C2 (f) E2: Photo taken at right side of participant seated at C2 (g) E2: Photo taken at left side of participant seated at C2

Figure 6. Pictures of E1 & E2 room taken from different perspective relative to the participant seated at different seat positions.

No.	Statement
1	Adviser was very friendly
2	I am please with adviser
3	I trusted adviser's advice
4	I enjoying playing with adviser
5	I would like to meet adviser face to face
6	Adviser gave good advice
7	Adviser was certain about the answer
8	I liked adviser
9	I relied mostly on adviser's advice

Table 1. Statements for post-experimental assessments of the adviser.

positions (see Figure 5). The legs for all those three isosceles triangles, which is the distance between participant and display, were maintained the same at 140 cm. We ensured that the vertical alignment of the eye level of viewers and the eye level of the avatar of the actor on the two displays were the same.

Incentives & risk

For both experiments, the number of chocolates that participants received was linked to the number of correctly answered questions. The number of chocolates varied between one and six.

Measurement Instruments

Task Performance Measure

The measure of advice seeking was defined as the proportion of one adviser being asked out of the total number of times advice was sought by a participant. As each participant had two advisers, but could only choose one of them for advice on each question, the following relationships hold: expert advice seeking = 1 – non-expert advice seeking, and one display advice seeking = 1 – the other display advice seeking.

Post-Questionnaire

Participants were presented with the post-experimental questionnaire with 9 statements (see Table 1) eliciting their subjective assessment of the two advisers, with 4 items measuring trustworthiness (Statement 3, 6, 7 & 9) and 5 items measuring enjoyment (Statement 1, 2, 4, 5 & 8). Agreement with the statements was elicited on 7-point Likert scales with the anchor 1 (Strongly disagree) - 7 (Strongly agree).

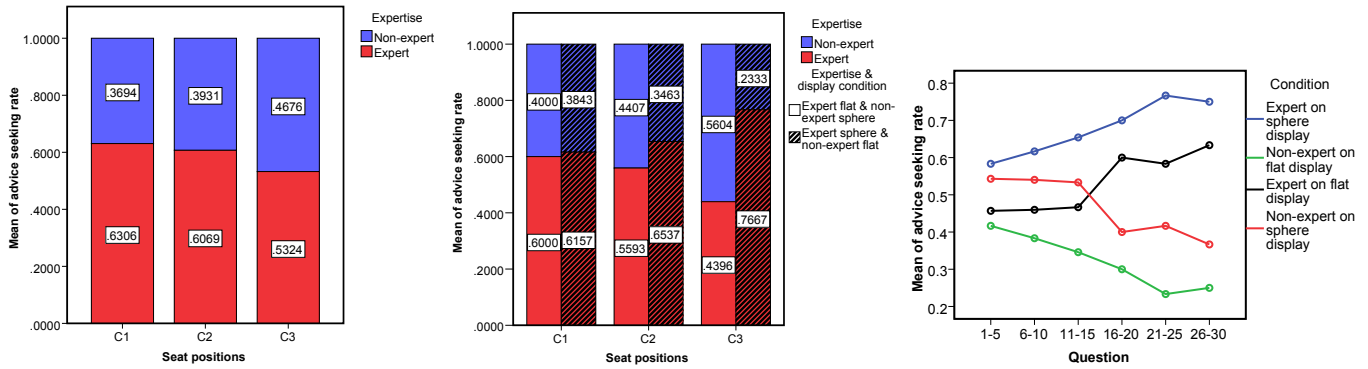
Open question

We asked each participant to write down his or her comments with a final open question: “For each round of games, please describe how you decided which adviser to rely on”. The purpose of this open question was to help explain some observed events during the game and to guide future research.

HYPOTHESES

Hypothesis 1: For E1 & E2, we expect participants to seek more advice from the expert adviser than the non-expert adviser. For E1, we expect that the more the seat position diverges from the central viewing position, the worse the observer will be able to discriminate between trustworthy and less trustworthy advisers. This is because the observer cannot look straight into the display and the slight visual spatial degradation will reduce observer’s ability to discriminate [14]. However, by introducing the spherical display (E2), we expect that the observer’s sensitivity to cues of trustworthiness to remain stable for all seat positions, as it conveys the same amount of information for all directions.

Hypothesis 2: For E2, we expect that the flat display will result in less trust compared to the sphere display. In other words, bias will occur when advice is preferred due to its display mode, irrespective of expertise. We further expect a negative bias towards the flat display representation will be found at off-center viewing angles, due to the loss of eye contact.



(a) E1: Advice seeking across seat positions. (b) E2: Advice seeking across seat positions. (c) E2: Advice seeking over time.

Figure 7. Results of E1 & E2: task performance measure.

EXPERIMENT 1

Method

Twelve participants (6 male), students and staff at University College London, were recruited to take part as observers in E1. The median age was 21.75 (SD = 3.20).

E1 had a one-way within-subjects design (see Figure 6(a) to Figure 6(c)). Every participant took part in the experiment at 3 different seat positions (C1, C2 and C3). The order of the answer options (A-D) of questions was randomized. The expertise and the participant’s 3 different seat positions order were counterbalanced, in order to reduce any confounding influence of the experiment environment such as lighting conditions and the orderings such as learning effects or fatigue.

Prior to starting the assessed part of the experiment, each participant completed a training round that consisted of easy questions where both advisers gave identical and correct advice. Then, participants answered 10 assessed questions each round. The participant could ask for advice before answering each question. For each question, participants could ask for advice from one of two advisers without knowing of adviser’s role of expertise. After each participant played one round at one seat position, the participant moved to another seat position. This process repeated for three different seat positions. Each participant has the same two advisers (Emma and Katy) for the whole study. After completing all rounds they were presented with the post-experimental questionnaire and an open question. Finally, the participants were compensated with chocolates based on their performance. The experiment took about 30 minutes.

Results

In E1, participants sought advice on 29.33 out of 30 questions (97.78%) over 3 rounds. 7 participants (58.33%) sought advice in every question. There was no cost associated with seeking advice. Figure 7(a) shows that the experts (red bar) were chosen more often than non-experts (blue bar) for all three seat positions. However, from seat position C1 to C3, the expert advice seeking rate dropped off whereas non-expert advice seeking rate increased. We interpret this to indicate the decrement of sensitivity for cues of expertise.

A one-way repeated measures ANOVA was conducted to compare the effect of the expert advice seeking rate in 3 seat positions (C1, C2 or C3) conditions. There was a significant effect of seat positions, $F(2, 22) = 6.356, p < .01$. Three paired samples t-tests were used to make post hoc comparisons between conditions. When we did three paired samples t-tests, we increased our chances of finding a significant result when one did not exist. Instead of using the value .05 to decide if we had reached statistical significance, we would instead use the value .017 (= .05/3) as the cut off. A first paired samples t-test indicated that there was no significant difference between C1 ($M = 63.06\%, SD = .063$) and C2 ($M = 60.69\%, SD = .091$) conditions; $t(11) = .945, p = .365$. A second paired samples t-test indicated that there was a significant difference for C1 and C3 ($M = 53.24\%, SD = .088$) conditions; $t(11) = 3.457, p = .005$. A third paired samples t-test indicated that there was no significant difference between C2 and C3 conditions; $t(11) = 2.304, p = .042$. The expert advice seeking rate at C3 is significantly less than C1. This suggests that further the seat position aside from the central position, the more difficulty the observer had in identifying the expert. This supports our first hypothesis.

EXPERIMENT 2

Method

Twenty-four participants (12 male) took part in E2. The median age was 21 (SD = 2.30). Participants had not previously interacted with advisers.

E2 is similar to E1, except that instead of presenting two advisers on two identical flat displays, we presented one on sphere display, and the other on flat display. E2 had a 2 display modes (*Expert is sphere display vs. expert is flat display*) \times 3 seat positions mixed design, resulting in 2 between-subject conditions with 12 participants each (see Figure 6(d) to Figure 6(g)). In each between-subject condition, two advisers were available. Depending on the factor display mode, either the sphere display or the flat display adviser gave expert advice, while the other gave non-expert advice. The two display positions (left-right) were counterbalanced by switching around the sphere and flat displays. To moderate the effect introduced by evaluating a novel type of display, we asked

each participant to complete a practice round prior to starting the assessed part of the experiment.

Results

Display type & viewing angle

In E2, participants sought advice on 29.21 out of 30 questions (97.36%) on average. 15 participants (62.5%) sought advice in every question. Figure 7(b) shows that the experts (red bar) were chosen more often than non-experts (blue bar) for all three seat positions. The overall expert advice seeking rate (expert on sphere display + expert on flat display) were 60.79%, 60.65% and 60.31% at the seat position C1, C2 and C3, respectively. This indicated that the overall expert advice seeking rate remained the same among three seat positions, which were approximately 20% higher than overall non-expert advice seeking rate. Figure 7(b) also shows that a preference for choosing sphere display increased from seat position C1 to C3, while decreased in the flat display condition. The overall sphere display seeking rate (expert on sphere + non-expert on sphere) were 50.78%, 54.72% and 66.35% at the seat position C1, C2 and C3, respectively. Sphere display advice seeking rate was higher in seat position C3. We note that for seat position C1, the flat display and sphere display were chosen equally often. This is expected as the seat position only slightly diverges from the front and the faces of two advisers can be seeing similarly on both display types.

A 2 (display: flat vs. sphere) \times 3 (seat positions: C1, C2 or C3) repeated measures ANOVA was conducted on the expert advice seeking rate, with display as a between-subjects factor and seat positions as a within-subjects factor. This revealed a significant main effect of display, $F(1, 22) = 13.757, p < .01$, indicating that expert advice seeking rate was significantly higher for sphere display. There was no significant main effect of seat positions, $F(2, 44) = .011, p > .05$, indicating overall expert advice seeking rate at different seat positions were not statistically significant different from one another, thus further supporting our first hypothesis. However, the display \times seat position interaction was significant, $F(2, 44) = 11.745, p < .001$, indicating that expert advice seeking rate due to seat position was presented differently in sphere and flat display conditions. This supports the second hypothesis.

We further investigated sphere display non-expert advice seeking rate at three different seat positions (unshaded blue bar in Figure 7(b)). The non-expert advice seeking rate $< .5$ would provide evidence for users' ability to discriminate between expert and non-expert advisers, whereas the value $> .5$ would be a sign of bias outweighing discrimination. Based on a one-sample t-test, the sphere display advice seeking rate at seat position C1 is significantly below .5, $t(11) = -2.582, p < .05$. There is also some indication at seat position C2, $t(11) = -2.111, p = .058$. However, no such effect is presented at seat position C3, $t(11) = 1.781, p > .05$, indicating that a bias towards sphere display is interfering with users' ability to discriminate, thus supporting the second hypothesis.

We then analyzed how participant's advice seeking behavior changes over time. Figure 7(c) presents the mean advice

seeking rate of every five questions in chronological order. The choice to seek advice from a specific adviser could be expected to depend upon the information accumulated from previous pieces of advice. It was thus assumed to be relatively arbitrary in first interactions. Participants increasingly sought advice from the expert as they gained experience with the advisers, but there is a bias towards sphere display. This gives us a further evidence for the second hypothesis.

Post-Questionnaire

Figure 8 shows the result of participant's self-reports. In the expert on sphere display condition, the statements measuring ability (Statement 3, 6, 7, & 9) were higher for the sphere display; and in the expert on flat display condition, those statements were higher for the flat display. This indicated that participants were able to identify the trustworthy adviser. However, statements measuring enjoyment (Statement 1, 2, 4, 5, & 8) showed similar or higher level of score for sphere display, even in the expert on flat display condition. This indicated that using the sphere display could increase social presence.

The responses to each statement item given by all the participants were averaged to create an aggregate response. We calculated Cronbach's alpha as the reliability test. The questionnaire measured four subscales: trustworthiness of the sphere display adviser (4 items, $\alpha = .893$), trustworthiness of the flat display adviser (4 items, $\alpha = .96$), enjoyment of the sphere display adviser (5 items, $\alpha = .807$), and enjoyment of the flat display adviser (5 items, $\alpha = .932$). We then analyzed the post-experimental assessments of the advisers by comparing each participant's rating of the sphere display adviser to that of the flat display adviser, irrespective of the expertise of each adviser. Significant differences in the post-experimental assessment (see Table 1) between sphere display and flat display adviser are thus indicators of bias on one subscale for one specific display type. Two paired-samples t-test were conducted to compare the ratings of the trustworthiness and enjoyment in sphere display and flat display conditions. Notable bias was found for enjoyment, sphere display rated as being friendlier than flat display, irrespective of expertise, $t(23) = -2.228, p < .05$.

DISCUSSION

We compared the advice-seeking rate at three seat positions between E1 and E2. E1 utilized two flat displays, with results demonstrating that participants' sensitivity to correct advice decreased at the far off-center viewing positions (C3). By introducing the spherical display in E2 that was able to preserve correct gaze direction and a simple pseudo-3D experience by providing perspective-correct rendering at all viewing angles using non-planar surface, we found participants' ability to discriminate remained stable at all viewing positions.

From participants' behavioral measures, we found that participants mostly chose expert advice in both flat and sphere display representations. This indicates that participants were able to discriminate between experts and non-experts, and accordingly, distributed more trust to the expert. However, there was also evidence that display representation can interfere with participants' ability to discriminate effectively. The sphere display produced a higher rate of advice seeking

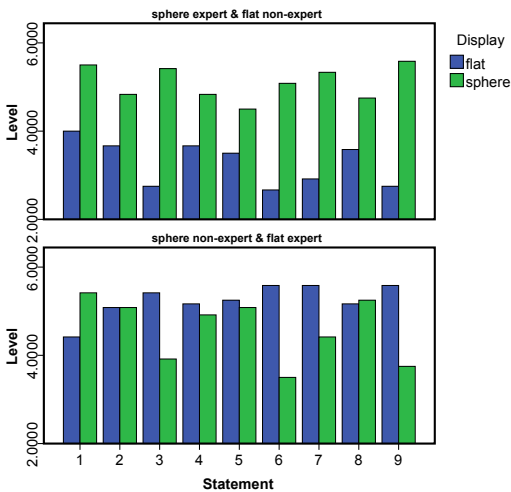


Figure 8. Post-experimental assessments of the advisers.

compared with the flat display. This behavioral manipulation emerged at off-center viewing positions and increased as the viewing position became more extreme. At the most extreme viewing position (C3), the rate of advice-seeking from the avatar displayed on the sphere was significantly greater than that sought from the avatar shown on the flat display. The preference for seeking advice from the avatar on the sphere display almost matched the preference for choosing expert advice, despite participants generally knowing on which display the expert was positioned. This negative bias towards the flat screen at off-center viewing angles in avatar-mediated communication parallels a similar finding by Nguyen et al. [19] in video-mediated communication. In that study, they examined the effects of spatial faithfulness on trust formation in a cooperative investment task. They found the spatial distortions of traditional flat display negatively affect trust formation patterns. The finding that trust can be easily and significantly manipulated in mediated interaction by adjusting display viewing angle has significant implications for telecommunication in general. We plan further investigation on this topic, with our next step being to quantitatively evaluate gaze of participants using eye tracking and introduce another between-subject condition (sphere display expert vs. sphere display non-expert) to further explore this finding.

Our post-experimental open question further supports our findings. Note that Katy was the adviser on flat display and Emma was the adviser on sphere display. In the expert is sphere condition, one participant stated “It is difficult to see Katy speak and look at her expressions while she answered, I could not feel good to communicate with Katy. Thus, I chose Emma more times.” Regarding viewing angle, another participant stated “I was sitting facing them directly rather than an angle with Emma, the more I felt they were reliable.” In the expert is flat condition, one participant expressed “Emma’s eye gives a supporting feeling, but Katy’s voice is more confident. Katy seems always certain about the answer, but Emma seems to tell what she knows.” Participants’ answers also show that there were other factors influencing their decision making, with one stating “I got a fully confident answer by

myself and Katy also told me the matched answer, so I tended to ask her more.”, and another stating “The longer time I spend with Katy and Emma, I figure out who knows more answers. But sometimes I still need to double check.”

CONCLUSION

We have presented a spherical display featuring a view-dependent rendering method to represent virtual avatars. We detail a method for enabling the displayed avatar to reproduce the facial expression captured from a person in real-time and with high-fidelity. The system provides observers with perspective-correct rendering and the nature of the display offers surrounding visibility whereas flat displays are only viewable from the front. This borderless spherical display can be statically situated as an interesting display for virtual avatars or other content. It could also be mounted on a robot as a mobile display for telepresence.

We investigated the display in the context of a trust scenario. We investigated the effects of display type (sphere and flat) and viewing angle for trust assessments in avatar-mediated interaction. While participants were able to discriminate trustworthy and less trustworthy advisers irrespective of display type, a negative bias for flat display can interfere with users’ ability to discriminate effectively. The interference became significant at off-center viewing angles, where the flat display no longer allows an undistorted and clear view. This demonstrates that a participant’s level of trust can be manipulated during avatar-mediated communication by the appearance of a remote interactant.

As the amount of our time spent in mediated interaction increases, these findings have significant implications for teleconferencing in general, and we plan to further investigate the phenomenon with classic video conferencing setups. The surrounding characteristics of spherical displays allow perspective-correct imagery to be seen from all viewing directions, and hence avoid the problems we have observed with traditional flat displays. By preserving a virtual avatar’s correct appearance and gaze direction, the spherical display is able to maintain a consistently high level of trust regardless of viewing position. Trust is a fundamental part of human interaction and we intend to explore other important scenarios and natural interaction in future work.

ACKNOWLEDGMENTS

The authors would like to acknowledge the FP7 EU collaborative project BEAMING (248620) and the China Scholarship Council. We thank the UCL VECG group members for their general support and the anonymous reviewers for their feedback.

REFERENCES

1. Andrist, S., Pejisa, T., Mutlu, B., and Gleicher, M. Designing effective gaze mechanisms for virtual agents. In *SIGCHI*, ACM (2012), 705–714.
2. Bailenson, J. N., Yee, N., Merget, D., and Schroeder, R. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in

- dyadic interaction. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 359–372.
3. Barnett White, T. Consumer trust and advice acceptance: The moderating roles of benevolence, expertise, and negative emotions. *JCP* 15, 2 (2005), 141–148.
 4. Bekkering, E., and Shim, J. Trust in videoconferencing. *Commun. ACM* 49, 7 (2006), 103–107.
 5. Bickmore, T., and Cassell, J. Relational agents: a model and implementation of building user trust. In *SIGCHI*, ACM (2001), 396–403.
 6. Bohannon, L. S., Herbert, A. M., Pelz, J. B., and Rantanen, E. M. Eye contact and video-mediated communication: A review. *Displays* (2012).
 7. Bos, N., Olson, J., Gergle, D., Olson, G., and Wright, Z. Effects of four computer-mediated communications channels on trust development. In *SIGCHI*, ACM (2002), 135–140.
 8. Chen, M. Leveraging the asymmetric sensitivity of eye contact for videoconference. In *SIGCHI*, ACM (2002).
 9. Dawes, R. M. Social dilemmas. *Annual review of psychology* 31, 1 (1980), 169–193.
 10. Desai, M., Tsui, K. M., Yanco, H. A., and Uhlik, C. Essential features of telepresence robots. In *TePRA*, IEEE (2011), 15–20.
 11. Deutsch, M. Trust and suspicion. *The Journal of conflict resolution* 2, 4 (1958), 265–279.
 12. Ducheneaut, N., Yee, N., Nickell, E., and Moore, R. J. Alone together?: exploring the social dynamics of massively multiplayer online games. In *SIGCHI*, ACM (2006), 407–416.
 13. Grossman, T., and Balakrishnan, R. The design and evaluation of selection techniques for 3d volumetric displays. In *UIST*, ACM (2006), 3–12.
 14. Horn, D. B., Karasik, L., and Olsen, J. S. The effects of spatial and temporal video distortion on lie detection performance. In *SIGCHI*, ACM (2002), 714–715.
 15. Kim, K., Bolton, J., Girouard, A., Cooperstock, J., and Vertegaal, R. Telehuman: effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod. In *SIGCHI*, ACM (2012), 2531–2540.
 16. Lincoln, P., Welch, G., Nashel, A., Ilie, A., Fuchs, H., et al. Animatronic shader lamps avatars. In *ISMAR*, IEEE (2009), 27–33.
 17. Mayer, R. C., Davis, J. H., and Schoorman, F. D. An integrative model of organizational trust. *Academy of management review* (1995), 709–734.
 18. Nguyen, D., and Canny, J. Multiview: spatially faithful group video conferencing. In *SIGCHI* (2005), 799–808.
 19. Nguyen, D. T., and Canny, J. Multiview: improving trust in group video conferencing through spatial faithfulness. In *SIGCHI*, ACM (2007), 1465–1474.
 20. Okada, K., Maeda, F., Ichikawaa, Y., and Matsushita, Y. Multiparty videoconferencing at virtual social distance: Majic design. In *CSCW* (1994), 385–393.
 21. Oyekoya, O., Steptoe, W., and Steed, A. Sphereavatar: a situated display to represent a remote collaborator. In *SIGCHI*, ACM (2012), 2551–2560.
 22. Pan, Y., and Steed, A. Preserving gaze direction in teleconferencing using a camera array and a spherical display. In *3DTV-CON, 2012*, IEEE (2012), 1–4.
 23. Parker, S. G., Bigler, J., Dietrich, A., Friedrich, H., Hoberock, J., Luebke, D., McAllister, D., McGuire, M., Morley, K., Robison, A., et al. Optix: A general purpose ray tracing engine. *TOG* 29, 4 (2010), 66.
 24. Rae, I., Takayama, L., and Mutlu, B. In-body experiences: Embodiment, control, and trust in robot-mediated communication. *interaction* 15, 28 (2013), 36.
 25. Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. The mechanics of trust: a framework for research and design. *IJHCS* 62, 3 (2005), 381–422.
 26. Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. Rich media, poor judgement? a study of media effects on users trust in expertise. In *People and Computers XIX The Bigger Picture*. Springer, 2006, 267–284.
 27. Roberts, D., Wolff, R., Rae, J., Steed, A., Aspin, R., McIntyre, M., Pena, A., Oyekoya, O., and Steptoe, W. Communicating eye-gaze across a distance: Comparing an eye-gaze enabled immersive collaborative virtual environment, aligned video conferencing, and being together. In *VR* (2009), 135–142.
 28. Sellen, A., Buxton, B., and Arnott, J. Using spatial cues to improve videoconferencing. In *SIGCHI* (1992), 651–652.
 29. Steptoe, W., Steed, A., Rovira, A., and Rae, J. Lie tracking: social presence, truth and deception in avatar-mediated telecommunication. In *SIGCHI*, ACM (2010), 1039–1048.
 30. Swol, L. M., and Snizek, J. A. Factors affecting the acceptance of expert advice. *British Journal of Social Psychology* 44, 3 (2005), 443–461.
 31. Teoh, C., Regenbrecht, H., and O’Hare, D. Investigating factors influencing trust in video-mediated communication. In *OzCHI*, ACM (2010), 312–319.
 32. Vertegaal, R., Weevers, I., Sohn, C., and Cheung, C. Gaze-2: conveying eye contact in group video conferencing using eye-controlled camera direction. In *SIGCHI* (Ft. Lauderdale, Florida, USA, 2003), 521–528.
 33. Weise, T., Bouaziz, S., Li, H., and Pauly, M. Realtime performance-based facial animation. *TOG* 30, 77 (2011), 1–10.
 34. Williams, E. Experimental comparisons of face-to-face and mediated communication: A review. *Psychological Bulletin* 84, 5 (1977), 963.
 35. Yamashita, N., Hirata, K., Aoyagi, S., Kuzuoka, H., and Harada, Y. Impact of seating positions on group video communication. In *CSCW*, ACM (2008), 177–186.