Eye contact and video-mediated communication: A review

Article *in* Displays · April 2013

DOI: 10.1016/j.displa.2012.10.009

4 authors, including:

Andrew Herbert
Rochester Institute of Technology

Jeff Pelz
Rochester Institute of Technology

Esa Rantanen
Rochester Institute of Technology

ARTICLE IN PRESS

Displays xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Displays

journal homepage: www.elsevier.com/locate/displa

Eye contact and video-mediated communication: A review

Leanne S. Bohannon [a,*], Andrew M. Herbert [b], Jeff B. Pelz [c], Esa M. Rantanen [d,1]

[a] Rochester Institute of Technology, 18 Lomb Memorial Drive, Rochester, NY 14623, United States
[b] Rochester Institute of Technology, Department of Psychology, 2323 Eastman Hall , 18 Lomb Memorial Drive, Rochester, NY 14623, United States
[c] Rochester Institute of Technology, Carlson Center for Imaging Science, 3112 Carlson Hall, 54 Lomb Memorial Drive, Rochester, NY 14623, United States
[d] Rochester Institute of Technology, Department of Psychology, 2353 Eastman Hall, 18 Lomb Memorial Drive, Rochester, NY 14623, United States

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Video-conferencing
Eye contact
Communication
Eye-tracking
Display
Laptop

ABSTRACT

A relatively new form of human communication, video-conferencing has become more popular as video technology improves and with increasing demands for real-time communication across greater distances. The full effects of video-conferencing on human communication are still being explored. Video-conferencing is presumed to be a somewhat richer form of communication than email and telephone, but not quite as informative as face-to-face communication. This review explores research into the influence of eye contact on communication and how video-conferencing mediates both verbal and non-verbal interactions. Facilitation of eye contact is a challenge that must be addressed so that video-conferencing can approach the rich interactions of face-to-face communication.

1. Introduction

In the history of human communication, video-conferencing is a relatively new means of correspondence dating from the introduction of the AT&T Picturephone at the New York World's Fair in 1964 [1]. However, the technology has advanced in leaps and bounds in the last 10 years, making videoconferencing available to anyone who can afford a modern laptop computer or a smartphone. In comparison to email, instant messaging and other electronically mediated communication, video-conferencing did not become commonplace until very recently. It is now standard for laptop computers and tablets to come equipped with a small camera mounted above the monitor, Apple's iPhones feature an application called FaceTime, Skype (video-conferencing software) has been purchased by Microsoft, and Cisco (a company that specializes in remote communication) has introduced ūmi, a video-conferencing system designed for home use. The recent revolution could also be driven in part by an increasingly global economy [2], family migration patterns, faster network connections, and substantial improvements in video-conferencing technology such as higher-resolution images and larger screens. Regardless of the causes for its growing popularity, like other technology-assisted communication media, video-conferencing reveals important aspects of human communication that should be considered in the development of new applications afforded by ever-advancing technology.

Along a continuum of communication media and their ability to manage rich information, or information that most reduces uncertainty and equivocation, face-to-face communication is the richest medium that provides immediate feedback and conveys many cues in the form of natural language [3]. Video-conferencing falls below face-to-face communication but above the telephone in terms of information richness. The importance of the view of the interlocutor's face is exemplified by the finding that including a still photograph to electronic mail can enhance the intended message and even prompt compliance with it [4]. Human communication is complex and based upon a combination of verbal and non-verbal cues to exchange information. Eye contact is arguably one of the most important non-verbal cues in communication, and we will examine its role in video-conferencing systems.

The unique properties of video-conferencing can impact our visual and verbal communication along with our perceptions of one another. Users of such systems also show substantial preference for video-conferencing that facilitates eye contact [5]. Understanding these effects can assist in system design and in mitigating obstacles to communication in further iterations of the technology. Ensuring natural (or the best approximation possible) eye contact in video-conferencing can increase the effectiveness of this mode of communication.

2. Eye contact in communication

Gaze can be defined as directed looking at any object, person, or direction. Eye contact is gaze directed at another's eyes, and mutual eye contact occurs when two people make eye contact simultaneously. When two people are in casual conversation, eye

* Corresponding author. Tel.: +1 585 210 8513.
E-mail addresses: Leanne.S.Bohannon@gmail.com (L.S. Bohannon), esa.rantanen@rit.edu (E.M. Rantanen).
[1] Tel.: +1 585 475 4412.

0141-9382/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.displa.2012.10.009

contact occurs about 61% of the time, about half of which is mutual eye contact [6]. People make eye contact 75% of the time while listening and 41% of the time while speaking [6]. However, these figures may vary based on the context of the conversation. While it seems clear that people spend more total time looking at the other while listening than speaking, the length of the gazes also vary. In an experiment with conversational dyads, participants made long gazes at their partner and looked away only briefly when listening, but made shorter gazes of equal length both at and away from their conversational partner while speaking [7]. Also, there is evidence that conversational partners tend to spend approximately the same amount of time looking at each other, adjusting their gaze based on the other's gaze [7]. The impact of gaze, and more specifically eye contact, on our social interactions as humans is significant. While we rely on a number of nonverbal cues to communicate information to others, eye contact seems to stand out distinctly from the rest.

2.1. Biological and developmental considerations

There is evidence that humans become sensitive to gaze direction by around 5 months of age [8], and infants as soon as 48 h after birth prefer faces that are gazing at them to faces that are looking away [9]. For example, infants looked significantly longer at photographs of faces whose eyes were looking straight ahead over faces with averted eyes. Furthermore, calm 9–12-week-old infants preferred a familiar face to a stranger's face when the familiar face was previously associated with both eye contact and sucrose delivery [10].

The amygdala, a brain structure related to processing emotional stimuli, is thought to be a key neural component of gaze detection; a patient with an amygdala lesion gazed more at people's mouths and less at their eyes during conversation as compared to subjects without an amygdala lesion [11]. Other cortical regions related to visual processing have also been shown to respond differentially for gaze direction. Area V5 of the visual cortex (part of the dorsal visual pathway involved in motion perception and coding actions) is more active when viewing images of faces whose eyes are animated to make eye contact as opposed to faces whose eyes are animated away from eye contact [12]. These results suggest that eye contact is an integral (or "hard-wired") part of human interaction.

2.2. Cultural, age, and sex differences

Eye contact during conversation varies from culture to culture, however. Arabs, Latin Americans, and Southern Europeans make more eye contact during conversation than Asians and Northern Europeans [6]. The English prefer to stand farther away and have unwavering eye contact while being attentive to what another is saying while Americans tend to stand closer and occasionally glance away from, and between, the two eyes of the person speaking [13].

In Japanese culture eye contact is considered rude, and people are taught to look at a person's Adam's apple instead of the eyes [14]. The level of eye-contact avoidance in the Japanese culture also depends on social rank – eye contact with a superior should most definitely be avoided [14]. The different effects of these cultural norms in communication are unknown. However, it is interesting to note that the lack of eye contact in one culture can be just as important a component of communication as the presence of eye contact in another.

In conversation, females make more eye contact while talking than do males [15]. The amount of gaze during conversation increases from between the ranges 4–6 and 6–9 years of age. Gaze decreases around 10–12 years of age, perhaps due to an increase in self-consciousness during those years of development, and increases again in adulthood [15].

2.3. Attention and memory

Gaze can be a powerful director of attention. To illustrate this, one study presented participants wearing an eye-tracker with a screen for 1500 ms displaying a fixation point between two squares [16]. Then the fixation point changed color. The color was a cue to the participant to look either to the left or the right. The fixation point disappeared after 150 ms, and then a photograph of a person's face appeared after either 50, 100, or 150 ms. The eyes of the face were looking either to the left, right, or center. The face stayed on the screen for 150 ms. In the trials with the 50 ms interval there were more antisaccades (saccades looking in the opposite direction of the color cue) when the eye direction was incongruent with the color cue. That is, the gaze direction of the face overrode the color cue instructions [16]. This is evidence that our attention tends to follow the gaze of another. Similarly, participants were faster to respond to a target when it was accompanied by a schematic face gazing in the same direction [17], possibly indicating that the tendency to follow another's gaze is reflexive in nature. Many illusions and card tricks work in part by misdirecting the gaze of viewers. Participants viewing a trick while being eye-tracked follow the gaze of the person performing the trick, thereby missing critical information that would reveal the trick [18]. The timing of gaze direction effects is consistent with prioritized visual processing of this information.

Memory, too, is affected by gaze direction. When presented with a prerecorded sales pitch for a soap product, participants remembered more about the product in the condition where the salesman made eye contact with the camera (the participant) than when he made no eye contact [19].

2.4. Impression formation

Gaze direction can influence the likeability and attractiveness of a person as perceived by another. When participants in a study [20,21] viewed female faces that made a gaze shift toward them, they rated the faces to be more likeable than when the eyes of the face made an animated shift away from the participant. Males rated female faces that made eye contact as more attractive than female faces whose eyes looked away [21]. Couples that scored higher on a questionnaire test measuring level of romantic love engaged in more eye contact than couples that had a lower romantic love score [22].

It also appears that believed eye contact, or what we are told about the level of our own or another's eye contact can affect our perception of that person [22]. Kleinke et al. [22] had male–female dyads engage in a conversation. After the conversation, the experimenter told the participants that the amount of gaze of one of them was either less, about the same as, or more than that of most people. Partners alleged to have made less than average eye contact were rated as less attentive. When males were told they gazed less than average at their female partner, they rated her more favorably. Conversely, females who were told they gazed at their male partner more than average rated him more positively. When males were told their female partner gazed an average amount, they rated her as more relaxed. Conversely, when females were told their male partner gazed an average amount they rated him as less relaxed. Males who were told their female partners had less than average gaze rated her as least attractive while females who were told their male partner had more than average gaze rated him least attractive. Both males and females who were told their partner had more than average gaze rated him or her as being more sincere [22]. Thus, regardless of the actual gaze of another

Please cite this article in press as: L.S. Bohannon et al., Eye contact and video-mediated communication: A review, Displays (2012), http://dx.doi.org/10.1016/j.displa.2012.10.009

person, the amount of gaze we *think* someone is directing at us affects our impression of him or her.

2.5. Other impacts of eye contact

Eye contact also plays an important role in other social interactions. For example, the presence or absence of eye contact can affect the perceived emotion expressed by the face. When presented with photographs of neutral faces, half of which gazed straight ahead (made eye contact) and half of which had averted gaze, participants were more likely to judge neutral faces that made eye contact as expressing anger or joy, whereas neutral faces with averted gaze were more likely to be rated fearful or sad [23]. When participants were presented with composite faces expressing both fear and anger, if the gaze was averted the face was more likely to be rated as fearful. When the composite faces were gazing at the participant, the face was equally likely to be rated as either fear or anger. Anger was rated as more intense when eye contact was present and fear was rated as more intense when gaze was averted [23].

In one unique experiment, an "invader" either stared or did not stare as he tried to "invade" the pew of local church-goers [24]. If the invader stared, the worshipper was likely to slide down the pew and make room for the person. However, in infrequent cases when the worshipper stared back at the invader, he or she typically did not make room for the invader. If the invader did not stare, the worshipper was less likely to move. It was uncertain if staring in this instance had to do with communicating aggression, or if its absence maintained a certain anonymity that made it easier for church-goers to ignore the request for a seat [24], but in any case underscores the importance of eye contact as nonverbal communication.

Eye-contact plays a role in perceived trust. US Customs inspectors were more likely to want to search a passenger in a mock interview when the passenger diverted his gaze frequently [25]. This was true despite the fact that gaze avoidance was not correlated with lying [25,26]. Eye contact impacts other social interactions such as impression formation, emotion and compliance and these affects will be examined later in this paper in the context of video-conferencing.

3. Eye contact in video-conferencing

Feeling as if the person with whom you are conversing is looking at your eyes is important; when given the option between using a video-conferencing system that enables eye contact and one that does not, people overwhelmingly prefer the system that enables eye contact [5]. The geometry of video-conferencing influences the perception of eye contact. Current video-conferencing technology is designed such that a camera sits atop a screen. The conversation partner's image appears on the screen and the camera captures one's face (and vice versa). For each conversation partner it appears as if the other person is looking downward – but each partner is looking right at his or her partner's eyes on the screen. This discrepancy between the camera's view of the head and where the person is gazing is a form of parallax. One would need to look directly into the camera for there to be eye contact. This is not what people do, because the intention is to look at the conversation partner's face, not the camera.

3.1. Cone of gaze

Early research on gaze perception focused on finding the threshold for what is considered eye contact. That is, how far can gaze deviate from the eyes and still be considered eye contact? Gamer and Hecht [27] coined the term "cone of gaze" to refer to the "range of considerable width wherein a person feels looked at" (p. 706).

Many studies examining gaze perception used an experimental setup involving pairs of people, one called the looker and the other called the observer [27–32] (see Table 1 for a summary). The looker is a confederate to the experiment and looks at different points on a scale invisible to the observer. The observer then makes some type of judgment about where the looker is looking. These gaze detection studies used different metrics for measuring the just noticeable deviation in gaze; some studies use degrees of visual angle and others use centimeters. When converted to degrees of visual angle the just noticeable deviation in gaze varies between 0.73° and 9.3°. This large variation between results is not surprising given the methodological differences in the studies. Thus, it is perhaps not useful to compare results directly to one another but instead to look for patterns; the important information to take from research on gaze direction perception is that (1) people are very sensitive to deviations in gaze from eye contact (2) except, perhaps to downward gaze and (3) perception of gaze direction is influenced by head orientation. All three of these points are relevant to video-conferencing.

3.2. Gaze perception during video-conferencing

Some studies on gaze perception focus specifically on eye contact perception and video-conferencing. These studies explore the role of video and parallax on gaze perception.

Chen [32] setup a "recording studio" in which a camera was placed behind a hole in the middle of a projector screen. Points were drawn 1° apart on the screen along a horizontal, vertical, and diagonal orientations indicating targets for the looker. The downward direction covered 15° and the other seven lines of points covered a range of 5° each to form a star-like pattern. The observer looked at videos of the looker gazing at the different points in random order. It was the observer's task to indicate whether the looker was or was not looking at him.

Observers were very sensitive to eye contact when the looker gazed up, left, or to the right but less sensitive when the looker gazed below the screen as shown in Table 1. When the looker was looking down, observers were much less sensitive to eye contact; the looker could look as far as 10° away before the observer reported eye contact was lost. One key difference between Chen's [32] study and a video-conferencing system is that participants in Chen's study made judgments about images of faces that were taken looking directly into the camera as opposed to from an elevated angle as is the case with video conferencing.

Chen [32] suggested that horizontal eye contact sensitivity to gaze deviation might be accurate because, as the eye rotates horizontally, the sclera becomes more visible. But as the eye rotates downward it becomes difficult to know if a person is making direct eye contact because less or no sclera can be seen. Chen [32] suggests we perceive downward gaze as great as 10° to be eye contact and recommends that in a video-conferencing system the eyes displayed on the screen should be located approximately 5° below the camera to create the illusion of eye contact.

To examine gaze direction perception under video-mediated conditions, Gale and Monk [33] had a looker and observer either sit across from each other with the targets between them, or a video camera was placed where the observer's head would have been. In both conditions both the looker and the observer were seated at eye-level with one another. Targets were marked on a vertical line that extended upward to the observer and on a horizontal arc such that every point on the arc was the same distance from the looker. Targets were marked such that the looker had to rotate his or her head 2° to look from one to the other. In one

Table 1
A summary of studies examining gaze perception sensitivity.

Study	Apparatus and stimuli	Angles	Distance	Task	Gaze sensitivity
Gibson and Pick (1963)	Targets marked on wall behind O	A 6 cm horizontal span with targets separated by 2.9°	200 cm	O responded "yes" or "no" to the question of whether or not L was looking directly at him	• The just noticeable deviation was 2.58° or about the distance from O's nose to the edge of her face
Cline (1967)	Transparent target board through which O viewed a mirror which reflected L	Targets along a vertical and horizontal axis separated by 2°	122 cm	O pointed to where she thought L was looking. She also said when she thought L was looking at her	• The threshold at which O still perceived L to be looking at him was: • 0.73° horizontal and 1.27° vertical for experiment 1 • 1.34° horizontal for experiment 2 and • 2.48° when O's head was rotated 30° right or left
Anstis et al. (1969)	Horizontal scale placed between L and O 6 cm above eye level. Scale rotated for experiment 2	Targets along a horizontal axis spanning 0°, 5°, 10°, 15°, 20° to the left and right	84 cm	O indicated on a scale where L was looking by pointing with a rod	• Overestimated where L was looking by between 50% and 87% • When L's head rotated 30° left of O, O estimated gaze to be 7.7° to the right. When rotated right, estimation was 5.44° to the left • When TV was rotated to O's left, estimation was 1.81° to the left. When rotated right then 4.56° to the right • When scale was vertical, estimates were true to scale
Chen (2002)	L's gaze positions were recorded by a camera through a hole in a projector screen which had targets drawn on it. L was prerecorded on video tape	Targets were separated by 1° increments downward to 15°. There were seven other lines radiating from the center that spanned 5° each and formed a star-like pattern	480 cm (O and L each sat 240 cm from camera/ screen)	O indicated if L was looking at her or not	• When L gazed up, down, left, or right, L could look at most 1° away from camera before O perceived L was no longer looking and her • L could look 10° downward from the center of the camera before O noticed L was no longer looking at her
Gamer and Hecht (2007)	Moveable eyes in a virtual and real looker eyes	Eyes of both the virtual and real head moved 1° at a time with the aid of computer software or an electronic pointer device, respectively. Eyes could move as far as 10° to the right or left	100 cm and 500 cm	O moved L's eyes either toward or away from him until he thought L was looking directly at him	• The cone of gaze was 9.3° at a distance of 100 cm and 8.2° at a distance of 500 cm for the virtual head • The cone was 8.2° at 100 cm and 3.9° at 500 cm for the real head • When L's head was rotated to right or left 10° the gaze was perceived to be shifted more in the direction of the head rotation than it actually was. The effect was greater for the 100 cm distance than the 500 cm distance

condition, observers were allowed to see the eye and head movements of the looker and in other conditions they were not. The looker first looked at the extreme points and center on the scale to "calibrate" the observer. The observer then made estimations about where the looker was looking.

The ability to see eye and head movements did not affect gaze direction perception ability. Gaze estimation was accurate to within about a target point (2°) for both live and video-mediated estimating conditions. Similar to Chen [32], gaze estimation error was smaller for horizontal than for vertical directions. In a second experiment the calibration was eliminated and observers made estimations about directions for both horizontal and vertical – gaze direction perception was also very accurate for the combination of directions [33].

Grayson and Monk [34] evaluated the perception of horizontal gaze direction with desktop video-conferencing systems. In one condition the camera was placed directly above the image of the looker and in the other it was offset. Additionally, the size of the looker's face on the screen was varied. The image was either large (352×288 pixels) or medium (176×144 pixels). The looker randomly looked (moving both eyes and head) at one of six points (43 mm apart) along a horizontal line that extended from either side of the image of the person on the screen. The observer estimated which point the looker was gazing at and the looker told the observer what the correct answer was.

Image size had no effect on the perception of gaze direction [34]. When the camera was positioned over the face of the participant, gaze estimation was quite accurate (84%) but when the camera was not place directly above the person being looked at, gaze estimation dropped (67% accuracy). When the camera is positioned directly over the image of the face, gaze direction perception accuracy decreases the further the gaze deviates from the centered "looking at me" position. Estimations showed little variation when the camera was offset.

In a second experiment, Grayson and Monk [34] examined gaze perception with a medium (176×144 pixels) and a small (88×72 pixels) image of the person on the screen. Also, in half of the trials the looker held a card with the fixation point on it over the image of the observer whenever he made a fixation – this was to eliminate the possibility that subtle communication could reveal gaze location when the looker was gazing at the image of the observer. Accurate detection of the eye contact position was 87% for the medium image and 68% for the small picture. Discrimination between points decreased for points further from the center. Making the observer invisible to the looker did not influence judgments of gaze direction.

Unlike most of the previous gaze research detailed earlier, these two studies [33,34] did not focus on thresholds of when perceived eye contact was lost. Instead, they focused on estimation of where someone was looking. When there is no parallax, estimation of gaze direction made from video seems to be quite accurate [33] but once parallax is introduced as in desktop video-conferencing, gaze direction perception accuracy decreases as the gaze direction deviates from a centered position, directly under the camera [34]. The varying effects of image size are consistent with the idea that small images lead to a poor user experience, but that camera position matters more than image size.

4. Effects of eye contact on video-mediated communication

4.1. Communication efficacy

Some researchers have used the *map task* as a tool to analyzing the effects of video-conferencing and eye contact on communication. The *map task* was originally suggested by Anderson et al. [35] as a way to teach and evaluate language. It is a useful tool to study communication because the *map task* requires information transfer and reception through spoken dialogue. In the *map task*, each participant is provided a map with landmarks scattered about it. The two maps contain some similar landmarks but are slightly different. One person is designated to be the instruction giver, and the other person is designated to be the instruction receiver. The instruction giver's map contains a trail. The goal of the task is for the giver to verbally articulate the trail's path to the receiver who must replicate the trail on his own map. Error in the final result can then be measured and provide one measure of success.

Boyle et al. [36] evaluated performance and communication efficacy on the map task in two conditions: co-present audio-only and face-to-face. Performance was measured by placing the instruction giver's map on top of the follower's map and calculating the difference in the paths (deviation) in square centimeters. While performance did not differ between the two conditions, communication patterns did.

Conversation efficiency was measured by number of turns, number of words, number of words per turn, number of interruptions (and interruption rate), and number of back channel responses (and back channel rate). Conversation pairs who could not see each other took more turns and used more words than those who could see each other. Participants who could not see each other took shorter turns, that is, they spoke fewer words per turn, than did participants who could see each other. Participants who could not see each other while completing the map mask interrupted each other twice as often as participants who could see each other. Conversations in which participants could not see each other were longer (as measured in number of words) than conversations in which participants could not see each other, and the interruption rate (percentage of turns containing interruptions was higher for the no-visibility condition. A back channel response was defined as "a turn which consisted of an 'uhuh' or 'mhm'" ([36], p. 10). Back channel responses were greater in both number and rate for conversations that enabled visibility than conversations that did not. Thus, the end product (the path) did not differ as a function of eye contact, but conversational efficiency was increased when the pair could see each other.

Boyle et al. [36] also examined the instruction follower's gaze toward the instruction giver. Conversations were videotaped and the transcribed dialogue was marked every time it was estimated the follower looked up and at the giver. When discussing a feature that differed between the two map versions, the instruction follower gazed at the giver more often than when discussing a feature that was the same. Other research examining gaze and turn taking found that people tend to look more at their conversational partner when they are ready to stop talking and let their partner begin speaking [7]. Participants in Boyle et al.'s [36] experiment may have looked at their partner more when discussing incongruent map features because they were trying to elicit verbal explanation from their partner regarding more details about the features.

According to Boyle et al.'s [36] measures of conversation efficiency, the ability to view one's conversation partner significantly increases the amount of information being conveyed in a shorter amount of time and with fewer words. It can be inferred that interruptions result from some sort of confusion about either whose turn it is or about the task. Back channel responses can be interpreted as a means of communicating understanding and agreement.

Doherty-Sneddon et al. [37] distinguished between the surface features and pragmatic function of dialogue by exploring dialogue and gaze under the context of five communication media – co-present audio, remote audio, face-to-face, video-mediated that enabled eye contact and video-mediated that did not enable eye contact. Like Boyle et al. [36], Doherty-Sneddon et al. [37] used the map mask as a cooperative task for their participants. Doherty-Sneddon et al. [37] analyzed the conversations between participants by use of Conversational Games Analysis. The Conversational Games Analysis technique, developed by Kowtko et al. [38], breaks the content of conversation into games that may contain several moves. Portions of conversations are given the analogy of a game because there are "rules" or conventions that are known by both parties.

In their first experiment, Doherty-Sneddon et al. [37] compared the dialogue of participants who did the paper and pencil map task under audio-only conditions to those who did the task face-to-face. In the audio-only condition, participant's dialogues contained 16% more conversational games than in the face-to-face condition. When participants could not see each other they had a greater need to confirm and check their understanding of one another; audio-only conditions contained significantly more conversational games that checked and confirmed information. Videos of the face-to-face conversation were coded for instances of gaze between participants. Doherty-Sneddon et al. [37] considered gaze to have occurred when they perceived one partner had looked at the face of the other. Participants tended to look at the other's face at points in the conversation when they were verbally eliciting feedback. Performance on the map task did not differ between conditions.

In their second experiment, Doherty-Sneddon et al. [37] compared dialogue structure and gaze between remote-audio (accomplished using microphones and speakers) and video-mediated communication that either enabled (by use of a video-tunnel, which uses mirrors to create the illusion of eye-contact) or prevented eye contact. Eye contact was prevented in one condition by raising the camera so that the participants were "now looking slightly down on his or her partner" ([37], p. 116). In contrast to Boyle et al. [36] who found that audio-only conversation required more words and turns, Doherty-Sneddon et al. [37] reported that there was no difference in number of words or turns between remote audio-only conversation and video-mediated conversation that did not enable eye contact. In fact, video-mediated conversation that enabled eye contact had significantly more words and turns than conversations in the no eye contact and audio-only conditions. Doherty-Sneddon et al. [37] explained this unexpected result by pointing out that the video-tunnel allowing eye contact may have been a novel experience for participants that fascinated them, resulting in more conversation. Another possibility is that the greater number of words recorded in the eye-contact condition could be explained by an increase in attention required by looking at another's face; some research suggests that increased eye contact during conversation increases the number of spoken hesitations. Repetitions, fillers like "er" and "um," false starts, and other remarks that may be irrelevant to semantic communication such as "I mean" and "sort of" were almost double when participants looked constantly at an interviewer as compared to when they were allowed to avert their gaze [39]. In both the video-mediated conditions, instruction givers initiated confirmed information significantly less than instruction givers in the audio-only condition. However, when video-conferencing was compared to remote audio-only conversations, there was no reduction of the number times instruction followers checked information. It appears as if, like face-to-face communication, video-mediated communication increases feedback to instruction givers when compared to audio-only communication. However, unlike face-to-face communication (in experiment 1), video-mediated communication does not reduce the need for instruction followers to check their knowledge of the instructions when compared to audio-only conversations. These findings suggest that, while video-mediated communication is an improvement over audio-only communication, it is not comparable to face-to-face communication in its ability to facilitate efficient communication.

Doherty-Sneddon et al. [37] analyzed gaze between the conditions and found that participants looked at each other significantly more in the video-conferencing condition that enabled eye contact than in the video-conferencing condition that did not enable eye contact. They also gazed more in the video-conferencing condition that enabled eye contact than in the face-to-face condition of the first experiment.

4.2. Trust

Besides establishing mutual knowledge, other important social interactions take place during communication and can be impacted by video-conferencing. Building trust, impression formation and establishing social status are all affected by different forms of mediated communication.

To determine the effect of eye contact in video-mediated communication on trust, Bekkering [40] created a scenario in which participants indicated the trustworthiness of a message delivered by people in three video-conferencing conditions. Participants were to imagine a scenario in which they had just been denied a job. The person at human resources informed them that the reason they did not get the job might have had something to do with a problem with letters of references. The participants were to imagine that they called the two people they had asked to write letters of references to check if letters had been submitted. Participants then listened to messages from the two people - one familiar person (their professor) and one unfamiliar person (a minister from a neighboring town) confirming that they did indeed send letters. The messages came in the form of an email, a voicemail, a video in which the person was making eye contact (the camera was eye level with the face), a video in which the camera was placed above the person's face, and a video in which the camera was placed to the side of the person's face. Each participant listened to the messages in all their forms (the script was the same for each) and indicated how much they trusted that the person had sent in their letter of reference.

Whether or not the person delivering the message was familiar to the participant did not affect perceived trust. When compared to the video that enabled eye contact (center camera), videos that prevented eye contact (side and top camera) resulted in lower perceived trust scores [40]. Trust scores did not differ between voicemail and centered video. Trust was significantly better for the centered video than for email. In summary, video-communication that prevents eye contact can reduce trust when compared to video-communication that does. Voicemail 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.

When participating in games that require group cooperation, participants trust each other more when communicating face-to-face as compared to computer-mediated media [41]. Participants, who played the Daytrader social dilemma game in groups of three, across four media conditions (face-to-face, video-conferencing, three-way phone conference, and text chat), had to make an investment choice that was either individual in which they were guaranteed two tokens for themselves or as a group in which all players were guaranteed three tokens so long as they all invested. Every five rounds of the game a bonus of 90 tokens was given to whoever had the most tokens at that point, and participants had the opportunity to discuss and agree upon a strategy for the next five rounds. If, for the five rounds preceding the bonus, all players cooperated and invested as a group, the bonus was split three ways.

Participants who played the game in the text chat condition made significantly fewer tokens from their investments than did groups in the other three conditions. Group investments started high and remained high for the face-to-face condition [41]. As the game progressed, group investments increased for the audio and video conditions. Group investments started low and stayed low for the text chat condition. If the amount of group investment is interpreted as trust, establishing trust was delayed in the audio and video conditions but gradually reached the same level as in face-to-face condition. Sometimes after each discussion, group members would gradually deviate from group cooperation. Cooperation was usually strongest right after a discussion. This measure of defection was used as an indicator of the fragility of trust [41]. Defection for the computer-mediated conditions was significantly greater than for the face-to-face condition. Bos et al. [41] interpreted these results to mean that trust is both more delayed and more fragile in computer-mediated-communication modes than in face-to-face communication.

Also using the Daytrader dilemma, Nguyen and Canny [42] examined trust under three media conditions: face-to-face, non-directional video-conferencing, and directional video-conferencing. Directional video-conferencing was defined as video-conferencing that enables accurate cues about gaze and gesture direction and was achieved by Nguyen and Canny's [42] MultiView video-conferencing system. This system differs from traditional video-conferencing systems in that a camera is placed in front of each member of the group giving them an accurate perspective of the other group members they are viewing on the screen.

The amount of cooperative investments made by groups did not differ between the face-to-face condition and the directional video-conferencing condition, and both of these were significantly higher than for non-directional video-conferencing. There was no difference in delay of trust between the three conditions. However, trust was significantly more fragile in the non-directional video-conferencing system than in both the directional video-conferencing system and the face-to-face conditions [42]. Additionally, participants' self-reported trust was significantly lower in the non-directional video-conferencing system than in both the directional video-conferencing system and the face-to-face conditions. The results of this study indicate 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 [42].

4.3. Impression formation

In a study particular to video-conferencing, Fullwood [43] used a unique interaction task to examine the role of communication media on impression formation. Same-sex participant pairs who were strangers to each other communicated either face-to-face or via video-conferencing. One partner in each pair was the "thought transmitter" and the other was the "mind-reader." The thought transmitter was given a card with a numbered list of 10 colors on it. The mind-reader was shown at list for 10 s that contained all possible colors he would be exposed to. The thought transmitters had to attempt to transmit each color to the mind-reader in the order that they appeared on the page by thinking about each color for 5 s. After each 5 s, the mind-reader had to guess what color the transmitter had been thinking of. The mind reader was not given feedback as to whether they had identified the colors correctly. The participants in the pair then swapped roles and afterwards filled out questionnaires about their partners which assessed their feelings about how much they liked the other participant, how intelligent they thought he or she was, and how good at reading minds they thought he or she was. The participants in the face-to-face condition liked their partners more than in the video-conferencing condition and rated them as more intelligent [43].

When groups of participants had discussions about a common strategy to survive on the moon (the Lost on the Moon Task) conditions enabling selective gaze had greater effects on behavior than other conditions [44]. Participants were more easily persuaded by a confederate to incorrectly rank order survival equipment when they were communicating over a video-conferencing setup that did not facilitate selective gaze compared to a video-conferencing setup that did enable selective gaze and via face-to-face communication [44]. Furthermore, when performing a murder mystery-solving task, groups of participants were more likely to communicate unshared information over video-conferencing than face-to-face [44].

A number of factors are likely responsible for the impact of video-conferencing on impression formation over face-to-face interactions. The quality and speed of the video and audio signal can mitigate the transmission of subtle non-verbal communication [45]. Furthermore, people may be more self-conscious when talking over video-conferencing (intensified by being able to see their own image on camera), and may therefore behave differently than they normally do [45]. Additionally, eye contact perception is not accurate in video-conferencing, and the presence or absence of eye contact can affect the impression of personality and emotions [45] [23].

5. Solutions to parallax in video-conferencing

The role of eye contact in human communication and the impact of the now "traditional" video-conferencing arrangement on perceived eye contact underline the importance of creating the optimal perception of eye contact during video-conferencing. Solutions to the lack of eye contact in video-conferencing are varied and involve hardware and software approaches. Perhaps one of the earliest solutions was the design of the video-tunnel in Doherty-Sneddon, et al. [37]. This comprises a half-silvered mirror placed at a 45° angle from a screen that is lying flat on the surface in front of the person communicating [46]. The image of the screen (and thus the other person's face) is reflected onto the mirror. A camera (or in some cases multiple cameras) is placed behind the mirror so that the person is both looking at the image of the eyes of the other and at the same time looking directly into the camera. This creates the appearance of eye contact. Such a setup can be limiting in terms of screen size and sometimes the person is required to place his head in a chin rest to maintain eye contact [46]. A modern-day attempt to adapt the video-tunnel to today's existing technology involves a miniature video tunnel that fits over small monitor-mounted web cameras [47].

Similar in concept to the video-tunnel is embedding a camera in the display. Tapia et al. [48] embedded a camera (1.5 mm in diameter) onto a surface and projected the other's face onto that surface. This enables the speaker to look at both the image of the other person and the camera, creating eye contact. This is not a viable solution for current laptops and monitors in commercial production.

Ott et al. [49] used a combination of multiple cameras and software to create eye contact. A camera was placed both above and below the screen; the top camera tilted slightly downward and the bottom camera tilted slightly upward. Together these cameras captured the same person from two different viewpoints. Corresponding pixels were merged and a combination image was used to create a single image in which it appeared the person was looking straight into the camera. However, it was not possible to capture the entire face resulting in some missing pieces in the final image [49].

While many alternate video-conferencing systems have been developed, the most prevalent video-mediated communication system remains the version where one camera is placed above the screen. Such systems are cost-efficient, portable, and compatible with many types of computer systems. Most laptops and tablets are now made with a small camera placed directly above the center of the display. Solutions that aim to improve eye contact

Please cite this article in press as: L.S. Bohannon et al., Eye contact and video-mediated communication: A review, Displays (2012), http://dx.doi.org/10.1016/j.displa.2012.10.009

perception under the constraints of this prevalent video-conferencing system are thereby most likely to be adopted by the majority of current users.

One such solution is software that alters the face image so as to create the illusion of eye contact [50]. The software replaces the eyes with computer-generated eyes and adjusts the head pose by mapping the face image onto a head model. The software is able to detect where the person is looking and adjust the appearance of the gaze accordingly. The disadvantage of the software is that head movements of more than 30° cause distortion.

However, simply by understanding human perception of gaze during video-conferencing, as was summarized in Section 4 of this paper, much simpler solutions are possible. One algorithm takes advantage of the large margin of error when perceiving gaze in a downward direction by locating the eyes of conversational partners and always displaying the image of the eyes close to the top of the display where the camera is located [51]. Without the use of extra equipment or computer programs, the average user can improve his or her video-conferencing experience by ensuring that his or her own image is close to the top of the camera's field of view, the image of the conversational partner is as close to the camera as possible, and by sitting back from the display until the perception of eye contact is achieved.

6. Conclusion

Much of the literature on gaze direction sensitivity emphasizes differences among samples (based on gender, culture, role, etc.). Perhaps the similarities should be emphasized. In general, allowing eye contact makes communication richer and more efficient at the same time. Trust is improved, with clear implications for how business transactions may improve with eye contact facilitated. There are very expensive solutions available for 'life-like' video-conferencing, but the vast majority of users are employing Skype or analogous free solutions on their existing computer systems. Based on the review of the literature, it seems clear that users can improve outcomes with a few simple measures: (1) by keeping the interlocutors' image as near the camera as possible; (2) by sitting back a bit to improve the *impression* of mutual eye contact; and, (3) smaller views of interlocutors should be avoided. The ordering of these three measures reflects their relative impact from most to least important given the literature reviewed. Video-conferencing software that facilitates such adjustments may provide better user experience. Furthermore, users can make communication via video-conferencing more like talking in-person by implementing some other simple methods - making sure hands and arms are in view of the camera assure that more non-verbal cues are communicated, and initially making face-to-face contact, or engaging in an ice-breaker activity builds report between communicants [45].

Research indicates that video-conferencing has an impact on verbal and non-verbal communication, also affecting social onstructs such as trust and impression formation. As video-conferencing becomes a more pervasive form of sharing information and relationships, technology needs to evolve to approach the effectiveness of face-to-face communication. Research on human sensitivity to gaze perception suggests the mounting of a camera above the screen is a good solution. Users appear to adapt to this viewing arrangement rapidly, but there is much greater satisfaction with systems such as directional video-conferencing or technical solutions such as the video tunnel or mounting a camera in the screen. Designers of video-conferencing systems need to be sensitive to the limitations of current systems, and research on gaze detection and eye contact needs to keep pace with technological innovations.

References

[1] AT&T, 1970: Picturephone, 2010. <http://www.corp.att.com/attlabs/reputation/timeline/70picture.html> (retrieved 28.10.10).
[2] L. Wolgemuth, Knights of the virtual table: video conferencing grows up thanks to 'telepresence' systems, US News and World Report 144 (2008) 48–50.
[3] R.L. Daft, R.H. Lengel, Organizational information requirements, media richness, and structural design, Management Science 32 (1986) 554–571.
[4] N. Guéguen, C. Jacob, Social presence reinforcement and computer-mediated communication: the effect of the solicitor's photography on compliance to a survey request made by e-mail, CyberPyschology and Behavior 5 (2002) 139–142.
[5] S.H. McNelley. The Significance of Eye Contact While Desktop Video Conferencing, Ph.D. Dissertation, United States International University, United States – California, 2001.
[6] M. Argyle, M. Cook, Gaze and Mutual Gaze, Cambridge University Press, London, 1976.
[7] A. Kendon, Some functions of gaze-direction in social interaction, Acta Psychologica 26 (1967) 22–63.
[8] A.J. Caron, R. Caron, J. Roberts, R. Brooks, Infant sensitivity to deviations in dynamic facial–vocal displays: the role of eye regard, Developmental Psychology 33 (1997) 802–813.
[9] T. Farroni, E. Menon, M.H. Johnson, Factors influencing newborns' preference for faces with eye contact, Journal of Experimental Child Psychology 95 (2006) 298–308.
[10] E.M. Blass, C.A. Camp, The ontogeny of face recognition: eye contact and sweet taste induce face preference in 9- and 12-week-old human infants, Developmental Psychology 37 (2001) 762–774.
[11] M.L. Spezio, P.S. Huang, F. Castelli, R. Adolphs, Amygdala damage impairs eye contact during conversations with real people, The Journal of Neuroscience 27 (2007) 3994–3997.
[12] S. Watanabe, R. Kakigi, K. Miki, A. Puce, Human MT/V5 activity on viewing eye gaze changes in others: a magnetoencephalographic study, Brain Research 1092 (2006) 152–160.
[13] E.T. Hall, The Hidden Dimension, Anchor Books Doubleday & Company, Inc., Garden City, NY, 1969.
[14] J.C. Nixon, J.F. West, Intercultural preparation for managers going to Japan, Mid-American Journal of Business 10 (1995) 57–64.
[15] M.H. Levine, B. Sutton-Smith, Effects of age, sex, and task on visual behavior during dyadic interaction, Developmental Psychology 9 (1973) 400–405.
[16] P. Ricciardelli, E. Bricolo, S.M. Aglioti, L.C. Chelazzi, My eyes want to look where your eyes are looking: exploring the tendency to imitate another individual's gaze, NeuroReport 13 (2002) 2259–2264.
[17] C. Friesen, C. Moore, A. Kingston, Does gaze direction really trigger a reflexive shift of spatial attention?, Brain and Cognition 57 (2004) 66–69
[18] G. Kuhn, B.W. Tatler, G.G. Cole, You look where I look! Effect of gaze cues on overt and covert attention in misdirection, Visual Cognition 17 (2009) 925–944.
[19] C. Fullwood, G. Doherty-Sneddon, Effect of gazing at the camera during a video link on recall, Applied Ergonomics 37 (2006) 167–175.
[20] M. Mason, E. Tatkow, C. Macrae, The look of love: gaze shifts and person perception, Psychological Science 16 (2005) 236–239.
[21] Z. Rubin, Measurement of romantic love, Journal of Personality and Social Psychology 16 (1970) 265–273.
[22] C.L. Kleinke, A.A. Bustos, F.B. Meeker, Effects of self-attributed and other-attributed gaze on interpersonal evaluations between males a females, Journal of Experimental Social Psychology 9 (1973) 154–163.
[23] R.B. Adams, R.E. Kleck, Effects of direct and averted gaze on the perception of facially communicated emotion, Emotion 5 (2005) 3–11.
[24] D.E. Campbell, G.E. Lancioni, The effects of staring and pew invasion in church settings, Journal of Social Psychology 108 (1979) 19–24.
[25] R.E. Kraut, D. Poe, Behavioral roots of person perception: the deception of judgments of customs inspectors and laymen, Journal of Personality and Social Psychology 39 (1980) 784–798.
[26] Communication. Merriam-Webster.com. (n.d.). <http://www.merriam-webster.com/dictionary/communication> (retrieved 02.08.12).
[27] M. Gamer, H. Hecht, Are you looking at me? Measuring the cone of gaze, Journal of Experimental Psychology 33 (2007) 705–715.
[28] S.R.H. Langton, H. Honeyman, E. Tessler, The influence of head contour and nose angle on the perception of eye-gaze direction, Perception and Psychophysics 66 (2004) 752–771.
[29] J.J. Gibson, A. Pick, Perception of another person's looking behavior, American Journal of Psychology 76 (1963) 386–394.
[30] M.G. Cline, The perception of where a person is looking, American Journal of Psychology 80 (1967) 41–51.
[31] S.M. Anstis, J.W. Mayhew, T. Morely, The perception of where a face or television "portrait" is looking, American Journal of Psychology 82 (1969) 474–489.
[32] M. Chen, Leveraging the asymmetric sensitivity of eye contact for videoconference, in: L. Terveen (Ed.), Proceedings of the AMC CHI 2002 Conference on Human Factors in Computing Systems, ACM Press, New York, 2002, pp. 49–56.
[33] C. Gale, A.F. Monk, Where am I looking? The accuracy of video-mediated gaze awareness, Perception and Psychophysics 62 (2000) 586–595.

Please cite this article in press as: L.S. Bohannon et al., Eye contact and video-mediated communication: A review, Displays (2012), http://dx.doi.org/10.1016/j.displa.2012.10.009

[34] D.M. Grayson, A.F. Monk, Are you looking at me? Eye contact and desktop video conferencing, ACM Transactions on Computer–Human Interaction 10 (2003) 221–242.
[35] A. Anderson, G. Brown, R. Shillcock, G. Yule, Teaching Talk; Strategies for Production and Assessment, Cambridge University Press, New York, 1984.
[36] E.A. Boyle, A.H. Anderson, A. Newlands, The effects of visibility on dialogue and performance in a cooperative problem solving task, Language and Speech 37 (1994) 1–20.
[37] G. Doherty-Sneddon, A. Anderson, C. O'Malley, S. Langton, S. Garrond, V. Bruce, Face-to-face and video-mediated communication: a comparison of dialogue structure and task performance, Journal of Experimental Psychology: Applied 3 (1997) 105–125.
[38] J.C. Kowtko, S.D. Isard, G.M. Doherty, Conversational Games Within Dialogue. Research Paper Published on the Human Communication Research Center, 1993. <http://www.hcrc.ed.ac.uk/maptask>.
[39] G.W. Beattie, A further investigation of the cognitive interference hypothesis of gaze patterns during conversation, British Journal of Social Psychology 20 (1981) 243–248.
[40] J.E. Bekkering, Visual angle in videoconferencing: the issue of trust, Ph.D. dissertation, Mississippi State University, United States – Mississippi, 2004.
[41] N. Bos, J. Olson, D. Gergle, G. Olson, Z. Wright, Effects of four computer-mediated communications channels on trust development, in: Proceedings of CHI 2002, ACM Press, New York, 2002, pp. 135–140.
[42] D. Nguyen, J. Canny, MultiView: improving trust in group video conferencing through spatial faithfulness, in: CHI 2007 Proceedings (Trust & Engagement), 2007, pp. 1465–1474.
[43] C. Fullwood, The effect of mediation on impression formation: a comparison of face-to-face and video-mediated conditions, Applied Ergonomics 38 (2007) 267–273.
[44] P.J. Werkhoven, J.M. Schraagen, P.A.J. Punte, Seeing is believing: communication performance under isotropic teleconferencing conditions, Displays 22 (2001) 37–149.
[45] C. Fullwood, M. Finn, Video-mediated communication and impression formation: an integrative review, in: Adam C. Rayler (Ed.), Videoconferencing: Technology, Impact and Applications, Nova Science, 2010, pp. 35–55.
[46] R. Vertegaal, I. Weevers, C. Sohn, C. Cheung, GAZE-2: Conveying Eye Contact in Group Video Conferencing Using Eye-Controlled Camera Direction, CHI 2003, ACM Press, 2003, pp. 521–528.
[47] See Eye2Eye: Teleprompter for Web Cameras, 2006. <http://www.bodelin.com/se2e/>.
[48] E.M. Tapia, S.S. Intille, J. Rebula, S. Stoddard, Concept and Partial Prototype Video: Ubiquitous Video Communication with the Perception of Eye Contact, Ubicomp 2003 Video, Program, 2003, 2 p. (un-numbered).
[49] M. Ott, J.P. Lewis, I. Cox, Teleconferencing eye contract using a virtual camera, in: S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, T. White (Eds.), INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems (Amsterdam, The Netherlands, April 24-29, 1993), CHI '93, ACM, New York, NY, 1993, pp. 109–110.
[50] J. Gemmell, K. Toyama, C.L. Zitnick, T. Kang, S. Seitz, Gaze Awareness for Video-Conferencing: A Software Approach. IEEE MultiMedia 7, 4 (October, 2000), 2000, pp. 26–35.
[51] J.W. McGowen, US Patent No. US8203595. Washington DC: US Patent and Trademark Office, 2012. <http://www.google.com/patents/US20100295920>.

Please cite this article in press as: L.S. Bohannon et al., Eye contact and video-mediated communication: A review, Displays (2012), http://dx.doi.org/10.1016/j.displa.2012.10.009