Human Interaction and Emerging Technologies (IHIET 2022), Vol. 68, 2022, 300–307
https://doi.org/10.54941/ahfe1002746

Gaze Behavior as an Objective Measure to Assess Social Presence During Immersive Mediated Communication

Ivo Stuldreher, Linsey Roijendijk, Maarten Michel, and Alexander Toet

Operations and Human Factors, Netherlands Organisation for Applied Scientific Research (TNO), Soesterberg, The Netherlands

ABSTRACT

Immersive communication systems provide increasingly realistic virtual environments, which may afford immersive social interactions that approach the quality of face-to-face (F2F) meetings by eliciting a sense of *social presence*; the feeling of being physically together with another person and having an affective and intellectual connection. To optimize a system's ability to convey social presence, there is a need for tools that efficiently and reliably measure the degree to which users experience social presence. Currently, the most widely used tools to measure (social) presence are questionnaires. As their ecological validity is questionable, there is a need for objective and non-intrusive measures to measure social presence during naturalistic social interactions. In our study, we aimed to identify a set of determinants of social presence that enable the assessment of a system's ability to convey social presence, preferably using easy to use, off-the-shelf tools. Considering eye gaze behavior is modulated by social presence and can be measured with relative ease for both F2F and mediated communication, we propose to use three eye gaze measures as an accessible means to assess the level of social presence a system can elicit.

Keywords: Social presence, Gaze behavior, Immersive mediated communication

INTRODUCTION

In our digital age, video- and audio-conferencing tools are increasingly used in human social interaction. New immersive communication systems based on virtual reality (VR), augmented reality (AR) or mixed reality (MR) may afford immersive social interactions that approach the quality of face-to-face (F2F) meetings. Examples include systems that enable photorealistic social VR communication (Gunkel et al., 2021) and simple AR systems that can be more easily applied in real-world environments (Abels et al., 2021). As with all immersive technologies, these systems enable a sense of spatial presence; the feeling of actually being in the virtual environment rather than in one own's physical environment (Biocca, Harms and Burgoon, 2003). However, immersive systems do not only aim to provide a feeling of being in the same (virtual) space with another person, but more importantly also to create an affective and intellectual connection with another person (Toet, Mioch, Niamut, van Erp, 2022). In other words, such systems aim for *social presence*.

As we will explain in this paper, our goal was not to further dissect the complex patterns that occur in social interaction or analyze which constructs predict social presence at any given moment during (mediated) communication, but to assess the degree of social presence through accessible means and equipment. We describe what social presence entails, how gaze behavior plays an essential role in social interaction, how gaze may reflect social presence, how gaze may be measured in F2F and virtual communication, and propose the next steps in the exploration of gaze as an objective measure to assess social presence during mediated communication.

ASSESSMENT OF SOCIAL PRESENCE

Social presence is a complex concept. It is often described as "the subjective experience of being present with a 'real' person" (Oh, Bailenson and Welch, 2018). To evaluate and optimize the ability to convey a high level of social presence in immersive systems, there is a need for tools that efficiently and reliably measure the degree to which users experience social presence. This requires an understanding of which aspects are linked with social presence. Social presence is divided into *copresence* – the feeling of being physically together with a communication partner – and *social interaction* – the feeling of having an affective and intellectual connection with one's communication partner (Toet et al., 2022). A systematic review by Oh, Bailenson and Welch (2018) provides a set of possible predictors for social presence. Some predictors are more related to the concept of copresence (e.g. physical proximity, depth cues), while others like visual representation, interactivity, agency and especially behavioral realism, focus more on social interaction aspects. We focus on the latter, with an emphasis on behavioral realism as this is at the core of social interaction, and always present in some degree in mediated communication.

Toet et al. (2022) recently described a holistic framework of social presence, in which the concept is broken down in five concepts based on a multisensory perception framework. Multisensory stimulation affects our brains at sensory, emotional, cognitive, behavioral, and decision-making levels. Toet et al. specify these five processing levels into five social presence quality factors: immediacy, intimacy, credibility, reasoning and (again) behavior. In their work, Toet et al. (2022) developed the Holistic Mediated Social Communication Questionnaire (H-MSC-Q) to capture social presence across these five quality factors.

Indeed, questionnaires are currently the most widely used tools to measure social presence. However, their ecological validity is questionable (De Moor et al., 2015), since they are intrusive and therefore typically applied after the experience itself, and require participants to report their own (often non-conscious) feelings and behavior. Preferably, the level of social presence would (also) be measured objectively without requiring a participant's subjective judgment. Such objective measures may be found in implicit behavioral and physiological responses.

However, Availability of implicit measures of social presence is limited. Spontaneous behavioral measures, such as startle responses, postural sway,

and conditioned social responses have been used as indicators of presence (Wiederhold, 2003), but mostly aimed at spatial instead of social presence. As behavioral measures are environment and content dependent, results are not easily generalizable to naturalistic interactions. Physiological parameters, such as brain potentials, electrodermal activity and heart rate are also described as potential indicators of presence (Grassini and Laumann, 2020), but no strong evidence for a link with social presence has yet been reported. We see most potential in eye gaze as an implicit and objective measure of social presence, since behavioral realism is a powerful predictor of social presence (Oh et al. 2018) and eye gaze behavior is an important part of social interaction, as we will show in the following paragraph.

GAZE BEHAVIOR IN SOCIAL INTERACTION

In natural – i.e. F2F – communication, eye gaze (where one looks, how long, and when) is an essential social interaction cue, especially when it is directed towards the face (Hessels, 2020). The movements, orientation, pupil size and blink frequency of the eyes convey important nonverbal signals that serve both interpersonal and practical functions. During social interaction, we focus on the eyes in an attempt to understand the other's intentions, beliefs, and emotions.

Direct eye contact however is only a small part of social interaction. We spend on average 60% directing our gaze towards the face, of which only 10% is directed at the eyes (Rogers et al., 2018) and these events last only around 2.2 seconds on average (0.36 seconds long for direct eye contact). Interestingly, people are not very sensitive to the gaze focus of one's partner upon one's face, and instead, generally perceive direct gaze towards their face as eye contact. There are indications that people look even less at the face (and eyes) during mediated communication and more at the background, compared to F2F contact (Freeth et al. 2013).

Gaze behavior also plays an important role in initiating and regulating social interaction. Speakers tend to look away from their partner as they begin talking to signal that they have the floor and that it is their turn to talk, and they look back when they are about to finish their turn in the conversation to signal that they are ready to hand control back to their partner (Kendon, 1967; Ho, Foulsham and Kingstone, 2015; Cañigueral and Hamilton, 2019). Dobre et al. (2021) found that people tended to have a higher frequency of gaze change (from averting to directing and vice versa) when they were being looked at, compared to when they were not. On group level, when being looked at the frequency of gaze change was between 0.54 and 0.60 changes per second, whereas when not being looked at the frequency of gaze change was between 0.42 and 0.51 changes per second. In a similar study, Freeth et al. (2013) found that participants changed their gaze more in a F2F conversation than when interact using mediated communication. In general, the number of changes of gaze-direction is highly correlated between partners in a dyad (Kendon, 1967).

Gaze orientation can both convey and direct attention (Frischen, Bayliss and Tipper, 2007). In conversations, gaze can be used to signal a person's level

of interest and attention (Bavelas, Coates and Johnson, 2002; Bavelas and Chovil, 2006) or degree of comprehension (Beebe, 1976). In collaborative spatial tasks, shared gaze can increase performance (Brennan et al., 2008; Yang et al., 2020). In sum, gaze behavior plays an important role in regulating a F2F social interaction.

GAZE BEHAVIOR AND SOCIAL PRESENCE IN CONTROLLED SITUATIONS

Existing work on mediated communication provides evidence that eye gaze behavior influences the level of social presence. However, most studies aim to find a general difference in the level of social presence by comparing extremes, rather than finding out when incremental changes in social presence occur. For instance, participants who had eye contact during a video conversation experienced more social presence than participants without any eye contact (Neureiter et al., 2013; Neureiter, Moser and Tscheligi, 2014). Similarly, Fauville et al. (2022) studied social presence using only screenshots of video conferences. They concluded that gaze direction has a medium-sized effect on social presence, suggesting the power of gaze in predicting the presence of a partner in the video conference (Fauville et al., 2022). Faces maintaining direct gaze (by looking straight into the camera) are rated as more socially present. Looking at the camera increased social presence, likeability judgments, and interpersonal attraction compared to looking at the screen or offscreen (Fauville et al., 2022).

Some studies do provide clues on how gaze behavior may differ with varying levels of social presence. Individuals look less at the eyes and heads of others and fixate on different facial areas when they believe they are talking to a live person (i.e., when they experience social presence) compared to a prerecorded video or computer generated agent (i.e., when they experience lower social presence) (Laidlaw et al., 2011; Gregory et al., 2015; Mansour and Kuhn, 2019; Holleman et al., 2020). The main factor for this reduction in gaze appears to be the belief that there can be real social interaction with the interlocutor (Gregory and Antolin, 2019). In two studies on mutual gaze, Bente et al. (2008) found that participants reported higher levels of social presence when their communication partner maintained longer mutual eye contact (2-4 seconds) with them compared to when he or she did not. However, when the mutual eye gaze was too long (8 seconds) to be behaviorally realistic, social presence was lower, which is in line with Roger et al.'s (2008) gaze duration values in F2F conversation. Gaze behavior during social interaction may therefore serve to assess the degree to which people experience social presence.

ESTIMATING GAZE BEHAVIOR

Technological advancements have made it possible to estimate gaze behavior in naturalistic environments and with little specialist equipment. Gaze estimation methods can be categorized into feature-based, model-based, and

appearance-based approaches (Hansen and Ji, 2010). Feature-based methods use eye features for gaze direction regression, such as corneal reflections caused by reflections of an external light source on the cornea. Model-based methods first detect visual features from a camera image, such as pupil, eyeball center and eye corners, and then fit a geometric 3D eyeball model to them to estimate gaze. Appearance-based methods also only require images obtained from an off-the-shelf camera, but directly learn a mapping from 2D input images to gaze directions using machine learning (Tan, Kriegman and Ahuja, 2002). Since there is no explicit eye feature detection step involved, this family of methods can typically handle input images with lower resolution and quality than model-based methods.

Gaze-based user modeling does not require specialized eye tracking equipment and gaze direction can be estimated for multiple persons in a single image. Machine learning technologies have significantly improved gaze estimation accuracy, enabling eye contact detection in natural multi-person interactions using off-the-shelf RGB (laptop) cameras (Zhang, Sugano and Bulling, 2017; Müller et al., 2018). Appearance-based methods are mainly of value for F2F scenarios, where traditional eye-tracking systems cannot be easily applied. However, for mediated communication using HMD's, the eyes are occluded, so an alternative method like eye tracking is required.

TOWARDS GAZE BEHAVIOR AS AN IMPLICIT MEASURE OF SOCIAL PRESENCE

We propose to deploy gaze estimation to automatically extract variables from gaze behavior that may reflect social presence during mediated communication. If proven suitable, gaze behavior provides a relatively easy measurement based on a single source and accessible equipment such as (often built-in) eye tracking and video footage. Although previous works suggest a role of gaze behavior on social presence, extracting information on social presence from gaze behavior is still relatively unexplored. Based on findings of gaze behavior in natural and mediated social interaction we see some promising parameters that may inform us on social presence.

Table 1 provides an overview of these parameters and the information required to assess these parameters. First, we hypothesize that if people are more socially present, they tend to look less at the eyes of their communication partner, but more at other facial areas. To be able to extract this information, one not only needs automatic gaze estimation, but also automatic face and eye detection in the gaze projection area. Second, we hypothesize that with increasing social presence, the frequency of gaze change from averting to directing the face will increase. Again one needs tools to automatically detect the face in the gaze projection area in addition to automatic gaze estimation. Third, we hypothesize that longer gazes to face (and eyes) will lead to increased social presence. However, if the gaze lasts too long (likely above 8 seconds (Bente et al., 2008)), the level of social presence will decrease. We propose to test these hypotheses in an experiment by comparing gaze behavior in F2F and mediated social interaction. Using a within-subject experimental design is advised to reduce variability in social responses. Besides automatic

Table 1. Gaze parameters that may be indicative of social presence.

Concept	Description & Hypothesis	Required Information
Eye vs. Face contact ratio	Communicating partners look less at the eyes of others and fixate more on different facial areas when they believe they are talking to a live person. We thus hypothesize, that with more social presence people tend to look less at the eyes of their communication partner, but more at other facial areas.	Automatic gaze estimation + automatic face and eye detection
Frequency of averting vs. directing gaze changes	People change their gaze direction from directing to averting the partner and vice versa more often when they are looked at. We thus hypothesize that with more social presence people tend to change their gaze direction more often.	Automatic gaze estimation + automatic face detection
Eye gaze duration	People look at their conversation partner only briefly each time, and prolonged eye gaze is perceived as unrealistic. We thus hypothesize that a longer mutual gaze results in higher social presence, however a persistent gaze from a (virtual) partner will decrease social presence.	Automatic gaze estimation + automatic face and eye detection

gaze estimation, one should also consider subjective self-report measures like the H-MSC-Q of Toet et al. (2022) in such an experiment. This would allow for a comparison of subjective and potential objective measures of social presence. In addition, it would allow for a coupling between eye gaze measures and the five quality factors of social presence and which factors should be addressed to improve the level of social presence of the (immersive) communication system. We would expect that our gaze measures correlate most with immediacy and behavioral factors, but how well eye gaze corresponds with intimacy, credibility and reasoning is not yet clear.

There are some caveats that should be considered when using gaze estimation to reflect social presence. First, gaze behavior is clearly not only mediated by social presence, but also by other factors such as attention (Frischen, Bayliss and Tipper, 2007) or the social context of the interaction (Hessels, 2020). In addition, there are cultural differences in mutual gaze during F2F interactions (Haensel, Smith and Senju, 2022) that make comparison across cultures difficult. Second, when comparing gaze behavior in F2F and mediated social interaction, different gaze estimation methods may be required. Whereas appearance-based gaze estimation is a suitable tool for F2F interactions, it is not applicable for mediated interaction when using head-mounted displays. When comparing gaze parameters between different interaction conditions with different gaze estimation methods, results may be influenced by consistent differences in accuracy and precision of gaze estimation.

REFERENCES

Abels, E.A.M. *et al.* (2021) 'Augmented reality-based remote family visits in nursing homes', in *IMX 2021 - Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*

Bailenson, J.N. *et al.* (2005) 'Transformed social interaction, augmented gaze, and social influence in immersive virtual environments', *Human Communication Research*

Bavelas, J.B. and Chovil, N. (2006) 'Nonverbal and verbal communication: Hand gestures and facial displays as part of language use in face-to-face dialogue', in *The SAGE Handbook of Nonverbal Communication*

Bavelas, J.B., Coates, L. and Johnson, T. (2002) 'Listener responses as a collaborative process: The role of gaze', *Journal of Communication*, 52(3).

Beebe, S.A. (1976) 'Effects of eye contact, posture and vocal inflection upon credibility and comprehension', *[Unpublished doctoral dissertation]*, pp. 1–35.

Bente, G., Rüggenberg, S., Krämer, N. C., and Eschenburg, F. (2008). Avatar-mediated networking: increasing social presence and interpersonal trust in net-based collaborations. Hum. Commun. Res. 34, 287–318.

Biocca F., Harms C., and Burgoon J. K. (2003) Toward a more robust theory and measure of social presence: Review and suggested criteria. Presence: Teleoperators and Virtual Environments, 12(5), 456–480.

Brennan, S.E. *et al.* (2008) 'Coordinating cognition: The costs and benefits of shared gaze during collaborative search', *Cognition*, 106(3).

Cañigueral, R. and Hamilton, A.F. d. C. (2019) 'The role of eye gaze during natural social interactions in typical and autistic people', *Frontiers in Psychology*, 10(MAR), pp. 1–18.

Dobre, G.C. *et al.* (2021) 'Direct Gaze Triggers Higher Frequency of Gaze Change: An Automatic Analysis of Dyads in Unstructured Conversation', *ICMI 2021 - Proceedings of the 2021 International Conference on Multimodal Interaction*, pp. 735–739.

Fauville, G. *et al.* (2022) 'Impression formation from video conference screenshots: The role of gaze, camera distance, and angle.', *Technology, Mind, and Behavior*, 3(1).

Freeth, M., Foulsham T. and Kingstone A (2013) What Affects Social Attention? Social Presence, Eye Contact and Autistic Traits. PLoS ONE 8(1): e53286.

Frischen, A., Bayliss, A.P. and Tipper, S.P. (2007) 'Gaze Cueing of Attention: Visual Attention, Social Cognition, and Individual Differences', *Psychological Bulletin*, 133(4).

Grassini, S. and Laumann, K. (2020) 'Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review', *Frontiers in Psychology*, 11.

Gregory, N.J. *et al.* (2015) 'Reduced gaze following and attention to heads when viewing a "live" social scene', *PLoS ONE*, 10(4), pp. 1–21.

Gregory, N.J. and Antolin, J. V. (2019) 'Does social presence or the potential for interaction reduce social gaze in online social scenarios? Introducing the "live lab" paradigm', *Quarterly Journal of Experimental Psychology*, 72(4), pp. 779–791.

Gunkel, S.N.B. *et al.* (2021) 'VRComm: An end-To-end web system for real-Time photorealistic social VR communication', in *MMSys 2021 - Proceedings of the 2021 Multimedia Systems Conference*

Haensel, J.X., Smith, T.J. and Senju, A. (2022) 'Cultural differences in mutual gaze during face-to-face interactions: A dual head-mounted eye-tracking study', *Visual Cognition*, 30(1–2), pp. 100–115.

Hansen, D.W. and Ji, Q. (2010) 'In the Eye of the Beholder: A Survey of Models for Eyes and Gaze', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), pp. 478–500.

Hessels, R.S. (2020) 'How does gaze to faces support face-to-face interaction? A review and perspective', *Psychonomic Bulletin and Review*, 27(5), pp. 856–881.

Ho, S., Foulsham, T. and Kingstone, A. (2015) 'Speaking and listening with the eyes: Gaze signaling during dyadic interactions', *PLoS ONE*, 10(8), pp. 1–18.

Holleman, G.A. *et al.* (2020) 'Implying social interaction and its influence on gaze behavior to the eyes', *PLoS ONE*, 15(2), pp. 1–27.

Kendon, A. (1967) 'Some functions of gaze-direction in social interaction', *Acta Psychologica*, 26(C), pp. 22–63.

Laidlaw, K.E.W. *et al.* (2011) 'Potential social interactions are important to social attention', *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), pp. 5548–5553.

Mansour, H. and Kuhn, G. (2019) 'Studying "natural" eye movements in an "unnatural" social environment: The influence of social activity, framing, and sub-clinical traits on gaze aversion', *Quarterly Journal of Experimental Psychology*, 72(8), pp. 1913–1925.

De Moor, K. *et al.* (2015) 'Quality of Experience: From Assessment to Application', *Report from Dagstuhl Seminar 15022*, 5(1).

Müller, P. *et al.* (2018) 'Robust eye contact detection in natural multi-person interactions using gaze and speaking behaviour', *Eye Tracking Research and Applications Symposium (ETRA)* [Preprint]

Neureiter, K. *et al.* (2013) 'Hands and Eyes: How Eye Contact is Linked to Gestures in Video Conferencing', *Conference on Human Factors in Computing Systems - Proceedings*, 2013-April(Vmc), pp. 127–132.

Neureiter, K., Moser, C. and Tscheligi, M. (2014) 'Look into my eyes & see, What you mean to me. Social presence as source for social capital', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8851, pp. 183–198.

Oh, C.S., Bailenson, J.N. and Welch, G.F. (2018) 'A systematic review of social presence: Definition, antecedents, and implications', *Frontiers Robotics AI*

Rogers, S. L., Speelman, C. P., Guidetti, O., and Longmuir, M. (2018). Using dual eye tracking to uncover personal gaze patterns during social interaction. Sci. Rep. 8, 1–9.

Tan, K.H., Kriegman, D.J. and Ahuja, N. (2002) 'Appearance-based eye gaze estimation', in *Proceedings of IEEE Workshop on Applications of Computer Vision*

Toet, A., Mioch, T., Gunkel, S.N.B. et al. (2022) Towards a multiscale QoE assessment of mediated social communication. *Qual User Exp 7, 4*.

Wiederhold, M.D. (2003) 'Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments, edited by Giuseppe Riva, Fabrizio Davide, and Wijnand Ijsselsteijn', *CyberPsychology & Behavior*, 6(6).

Yang, J. *et al.* (2020) 'The effects of spatial auditory and visual cues on mixed reality remote collaboration', *Journal on Multimodal User Interfaces*, 14(4).

Zhang, X., Sugano, Y. and Bulling, A. (2017) 'Everyday eye contact detection using unsupervised gaze target discovery', *UIST 2017 - Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 193–203.