The Relative Importance of Social Cues in Immersive Mediated Communication

Navya N. Sharan, Alexander Toet[(✉)], Tina Mioch, Omar Niamut, and Jan B. F. van Erp

Netherlands Organisation for Applied Scientific Research TNO, The Hague, Netherlands
{lex.toet,tina.mioch,omar.niamut,jan.vanerp}@tno.nl

Abstract. Effective interpersonal communication is important to maintain relationships and build trust, empathy, and confidence. In this digital age, communication has become mediated, which filters out many of the social cues that are essential to facilitate interpersonal communication. This paper investigates the extent to which social cues influence social presence in mediated, bidirectional, multiparty interaction. Literature related to six social cues – paralinguistic cues, linguistic cues, body language, eye movements, facial expressions, and proxemic cues – was reviewed. These cues were ranked based on how relevant they are in creating a sense of social presence in mediated social communication (MSC). The most relevant cue was eye movements, followed by facial expression and linguistic cues, and lastly, body language and proxemic cues. Paralinguistic cues could not be ranked due to sparse literature in the context of MSC. Further research is required to better understand how social cues can be incorporated into MSC systems.

Keywords: Social cues · Social presence · Virtual reality · Social VR · Remote communication

1 Introduction

We, as humans, have a basic need to communicate with others to maintain interpersonal relationships [1]. Effective interpersonal communication relies on more than only an efficient exchange of semantic content. Emotions, which form a large part of interpersonal communication [1], are mainly communicated through nonverbal behavior like facial expressions, intonation, touch, and gestures [2]. In today's digital age, this communication has become increasingly mediated. With the rise of computer graphics and communication technologies, we are getting closer to bringing mediated communication at par with real-life communication in the visual modality.

To allow mediated social communication (MSC) to afford the same affective characteristics as face-to-face (F2F) communication, MSC should be able to elicit a sense of social presence [3]. Social presence is defined by two key elements: (1) spatial presence and (2) social interaction [3]. Therefore, social presence is the sense of being together and interacting with other remote individuals in a mediated environment [4]. Social

presence is determined by the quality of direct mutual interaction [5], contextual cues, the telepresence (i.e., the feeling of being located in the mediated environment) qualities of the MSC, and individual user traits [6].

Contextual or social cues refer to verbal and nonverbal signals expressed through voice, face, body, or motion, and guide communication and other social interactions by influencing how we perceive and respond to one another. Co-presence, i.e. the sensation of being present in the same environment as other remote individuals in MSC, increases when individuals exhibit social interaction cues, even those as minimal as eye contact [7]. This, by extension, increases social presence. Social cues are diverse and complex but the ones relevant here are: (1) paralinguistic cues, (2) linguistic cues, (3) body language, (4) eye movements, (5) facial expressions, and (6) proxemic cues. This paper focuses on social cues perceived through vision and audition as opposed to olfaction and touch as the former are easily integrated into existing MSC systems, thereby generating a more extensive body of research.

Experiencing social presence in MSC involves a sense of intimacy (i.e., feeling connected with others during an interaction) and a sense of immediacy (i.e., the psychological distance between communication partners) [4]. Therefore, social presence is inherently bidirectional (i.e., involving a sense of mutual awareness) [8]. As such, this paper aims to answer the research question: *To what extent do social cues influence social presence in mediated, bidirectional multi-party interaction*? By reviewing current literature, first, the importance of social cues in facilitating social presence will be established. Next, the six identified social cues and how they influence mediated social presence will be explored. Finally, these cues will be ranked in terms of their effect on facilitating social presence in MSC.

2   Social Cues and Social Presence

Users of extended reality (XR, i.e., augmented reality, augmented virtuality, or mixed reality) based communication systems have the impression that their remote communication partners are physically present in their own real-world environment while users of immersive virtual reality (VR) communication systems experience being physically transported to another environment [9]. These systems can reproduce a wide range of social cues (visual, auditory, and sometimes haptic) and thus, improve social communication. Contrastingly, traditional communication tools like video- or audio-conferencing only transmit voice and non-verbal sounds, and partial two-dimensional images of communication partners, thereby filtering out many of the social cues essential for building trust, empathy, and confidence [3].

Additionally, the inability to transmit nonverbal cues can have a negative effect on interpersonal communication [4]. The presence of relevant social cues, in turn, increases social presence. Therefore, relevant social cues and other information about the shared environment need to be reliably conveyed in MSC systems. Nonverbal social cues create a sense of immediacy while conversations without such cues result in a "distant" interaction [10]. In other words, to provide communication partners a mutual understanding of each other's intentions, mood, and feelings, it could be beneficial to unambiguously represent social cues in MSC [11].

The following sections will address the extent to which the six identified social cues influence social presence in MSC. These cues have been broadly classified into verbal, nonverbal, and other cues.

3   Verbal Cues

3.1   Paralinguistic Cues

Paralinguistic cues are social cues that are embedded in vocal communication but are separate from actual language or semantic content. This includes vocal pitch, tone, loudness, and inflection. In digital textual communication, paralinguistic cues could also include emoticons [12] or reactions like "likes" or "dislikes". However, for the purpose of this paper, these will not be considered.

Literature on the relationship between paralinguistic cues and social presence is sparse. However, one study [13] found strong evidence suggesting that voices that sound like an extrovert induce a greater sense of social presence than voices that sound like an introvert. Additionally, social presence increased when users heard voices that manifested a personality similar to their own. This indicates that while paralinguistic cues influence social presence, further investigation is required to understand their role in modern MSC.

Paralinguistic cues are present in most MSC systems, including the simplest ones like audio-conferencing. Thus, these are the most ubiquitous of all social cues. A better understanding of the relationship between paralinguistic cues and social presence could have important implications for the development of all types of MSC systems.

3.2   Linguistic Cues

Linguistic cues refer to the choice of words and structure of sentences, i.e. dialect and syntax, that users employ during communication. One study [14] investigated how verbal communication influences social presence in a mixed reality remote collaboration task. Participants completed a search task in a virtual large office. The researchers classified verbal communication between participants into four types of utterances: (1) social/emotional expressions (i.e., expressions of feeling), (2) acknowledgement (i.e., indication of listening), (3) complete reference (i.e., statements that could be understood on their own), and (4) reference pronouns (i.e., statements that used pronouns to refer to things).

In conditions with only verbal cues, participants asked for clarification whereas in conditions with verbal and visual cues (e.g., gestures), conversation patterns were more natural. Participants were also more confident and expressed positive feelings. These findings illustrate that while linguistic social cues are essential for MSC, they are far more effective when paired with other nonverbal social cues.

4   Nonverbal Cues

4.1   Body Language

Body language consists of two key elements: (1) body posture and movements and (2) head movements and hand gestures. While the former is subtle, less definite, and can

indicate some feelings and attitudes, the latter is deliberate and important to communicate meaning without words. Some examples of body posture movements are arm- or leg-crossing. Head and hand gestures could include movements like nodding, pointing, waving, etc.

In the context of MSC, the extent to which visual representations of communication partners behave like an actual person (i.e., behavioral realism) influences social presence [6]. Here, visual representations include but are not limited to two-dimensional images, avatars, etc. The positive effects of behavioral realism are stronger when the visual representation's behavior indicates awareness of its communication partner's presence. A research [15] found that participants felt increased levels of social presence when they interacted with a computerized agent (i.e., not a visual representation of another human user) that nodded at appropriate times than an agent that did not. While interacting with an agent is different from communicating with other humans, this research emphasizes that participants who thought they were interacting with an avatar representation of another human did not behave differently from those interacting with an agent. Behavioral realism influenced both sets of participants.

In general, users evaluate gestural solutions positively. Research found that users are more involved, aware, and socially present when hand tracking was used to create meaningful gestures [16]. Moreover, body and hand gestures were preferred over keyboard and mouse interactions. These findings show that behavioral realism in body language, especially with head and hand movements, is a strong predictor of social presence.

4.2   Eye Movements

Eyes are an important aspect of nonverbal communication. The morphology of the human eye has developed to enhance gaze perception [17]. The dark iris within a large, white sclera creates a high contrast, allowing people to follow each other's gaze [18]. Eye movements include looking, staring, and blinking. People also use eye contact to build trust; a steady gaze is usually perceived as a sign of honesty whereas a shifty gaze and an inability to maintain eye contact is seen as an indicator of deception.

Eye movements in MSC have also been widely investigated. In a study by [19], participants listened to a presenter in an immersive collaborative virtual environment. Eye gaze was transformed strategically to either be (1) reduced (i.e., presenter did not face either of the two listeners) or (2) augmented (i.e., presenter continuously looked at both listeners). In the augmented condition, each listener felt that the presenter was only looking at them. It was predicted that, due to increased engagement and social presence, the presenter would be better able to persuade the listener in the augmented gaze condition. This hypothesis was supported for female participants only. On an explicit level, female participants reported the problematic aspects of the augmented gaze like the unresponsiveness and unawareness of the presenter, and this also reflected in the self-reported social presence ratings. However, on an implicit level, they were still persuaded by the visual input of the augmented gaze.

These results suggest that though users recognize the issues of artificiality associated with transformed gaze, "personalized" gaze still has the intended impact in terms of inducing a sense of engagement and social presence. This study further highlights the extent to which gaze direction, on its own, can influence social presence in MSC.

4.3 Facial Expressions

The first social cue that we typically perceive is a person's facial expression. Facial expressions form a huge proportion of nonverbal communication. While other nonverbal social cues can differ among individuals, facial expressions conveying certain emotions like happiness, sadness, anger, and fear, are largely universal.

Eyebrow movements are a part of many facial expressions and are probably the most relevant of all facial gestures in conversations. For example, literature shows that eyebrow movements correlate with speech and happen during pauses in speech or to emphasize words or phrases [20]. Since eyebrow movements happen during both speech and pauses, they are not directly linked but rather the result of linguistic and conversational choices. At the same time, eyebrow movements of listeners reassure the speaker that they have their listener's attention. Eyebrow movements play an important role in maintaining attention, thereby facilitating a sense of social presence.

Similarly, research [21] also found that participants felt an increased sense of social presence when a virtual agent blushed after making a mistake during a presentation as this reflects behavioral realism.

5 Other Cues

5.1 Proxemic Cues

Proxemic cues refer to how we perceive the space around us. These cues are only generated when there are two or more parties involved in communication. Personal space is also a nonverbal social cue, which depends on factors like sociocultural norms, situational factors, individual traits, and familiarity.

A study investigating human-robot interaction [22] found that proxemic behavior exhibited by robots affected how socially present they were perceived to be. Specifically, participants considered robots that acted passively as more socially present than those who acted assertively. For instance, passive proxemic robots gave "right of way" in a path-crossing event, which could have led participants to perceive these robots as attending to them and their behavior and as considering the social rules of politeness. Thus, allowing users to establish and respect personal space in MSC could increase social presence. However, there are some issues associated with generalizing findings with non-human entities to human-human MSC. Some studies [15] indicate that the belief of interacting with either humans or non-human entities lead to no substantial differences in expectations or behavioral reactions. This suggests limited generalizability, but further research is recommended.

Likewise, [14] also found that while completing a search task, the remote user's spatialized voice guidance improved the local user's performance and sense of social presence in a highly cluttered virtual environment by providing auditory clues for spatial directions. These results further support the idea that proxemic cues like personal space and spatialized audio improve social presence by keeping the user engaged with the MSC environment.

6   Ranking of Relevance

Based on literature, it is evident that some social cues are more relevant to facilitating social presence than others. This section outlines a ranking of relevance for the social cues highlighted here and how this could be incorporated into the design of MSC systems. Since there is limited research comparing different cues or ranking the cues, this section is reflective of the interpretation of existing literature.

As seen in the previous sections, the most pervasive, and thus, relevant social cue for MSC is eye movements, specifically gaze direction. Therefore, to improve social presence in MSC systems, communication partners should be able to make direct eye contact and see each other's gaze direction.

Facial expressions are the next most relevant social cue, but they frequently correlate with linguistic cues. Thus, to further increase social presence, communication partners should be able to observe one another's facial expressions synchronized with their linguistic social cues.

Next, behavioral realism in body language and proxemic behavior was also essential for social presence. While it is more difficult compared to the previous cues to incorporate these into MSC, allowing communication partners to perceive each other's body language and proxemic behavior could exponentially increase social presence.

Finally, for paralinguistic social cues, while it does, to some extent, influence social presence, further research is required to specifically establish its importance in facilitating social presence.

7   Further Research

There still remain several questions about social cues in MSC that lie outside the scope of this paper. To better understand how social cues should be incorporated into the design of MSC systems, further research along the following avenues is required.

Firstly, there could be cultural differences in terms of which social cues are relevant and the way in which they are interpreted. Currently, most research related to social cues and social presence is carried out with users from Western cultures. Research with users from non-Western cultures is required to design MSC systems for those users that use different social cues to achieve social presence.

Secondly, social behavior like the display of social cues is determined by individual traits [6]. Designing MSC systems which disclose more social cues than current systems already disclose could influence the social behavior of users. For instance, some users who are accustomed to audio-conferencing might find it difficult to use MSC systems that project their body language and eye movements in a mediated environment.

From a technological design perspective, it is unclear how incorporating more social cues will influence the design of the MSC system itself. For instance, the system might become heavier or the user interface might become too complex, thereby deterring usage. Further research is required to carefully design MSC systems that incorporate multisensory social cues.

Finally, the interplay between different social cues needs to be studied further. Social cues could have an interaction effect. This was briefly addressed in the previous sections. For instance, [23] found that speech included turn-taking and back-channel signals,

which were given through eye contact and head nods, respectively. This indicates that linguistic cues, eye movements, and body language often interact in social communication. Deeper research is required to understand how such interactions could influence the design of MSC systems.

8 Conclusion

This paper identified and reviewed literature related to six key social cues: paralinguistic cues, linguistic cues, body language, eye movements, facial expressions, and proxemic cues. These cues were ranked based on the extent to which they are relevant in creating a sense of social presence in MSC. The most relevant cue was found to be eye movements, followed by facial expression and linguistic cues, and thirdly, body language and proxemic cues. Paralinguistic cues could not be ranked due to sparse literature in the context of MSC. Avenues for further research are recommended to better understand how social cues can be incorporated into MSC systems.

References

1. Kjeldskov, J., et al.: Using cultural probes to explore mediated intimacy. Australas. J. Inf. Syst. **11**, F2008 (2004)
2. Knapp, M.L., Hall, J.A.: Nonverbal Communication in Human Interaction, 7th edn. Wadsworth, CENGAGE Learning, Boston (2010)
3. Toet, A., Mioch, T., Gunkel, S.N.B., Sallaberry, C., van Erp, J.B.F., Niamut, O.: Holistic quality assessment of mediated immersive multisensory social communication. In: Bourdot, P., Interrante, V., Kopper, R., Olivier, A.H., Saito, H., Zachmann, G. (eds.) Virtual Reality and Augmented Reality. EuroVR 2020, LNCS, vol. 12499, pp. 209–215. Springer, Heidelberg (2020)
4. Short, J., Williams, E., Christie, B.: The social psychology of telecommunications. Wiley, New York (1976)
5. Biocca, F., Harms, C.: Defining and measuring social presence: contribution to the Networked Minds theory and measure. In: Gouveia, F.R., Biocca, F. (eds.) Fifth Annual International Workshop on Presence, pp. 7–36 (2002)
6. Oh, C.S., Bailenson, J.N., Welch, G.F.: A systematic review of social presence: definition, antecedents, and implications. Frontiers Rob. AI **5** (2018)
7. Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M.: Interpersonal distance in immersive virtual environments. Pers. Soc. Psychol. Bull. **29**, 819–833 (2003)
8. Biocca, F., Harms, C., Burgoon, J.K.: Toward a more robust theory and measure of social presence: Review and suggested criteria. Presence Teleoperators Virtual Environ. **12**, 456–480 (2003)
9. Hartmann, T., Wirth, W., Vorderer, P., Klimmt, C., Schramm, H., Böcking, S.: Spatial presence theory: state of the art and challenges ahead. In: Lombard, M., Biocca, F., Freeman, J., Ijsselsteijn, W., Schaevitz, R.J. (eds.) Immersed in Media: Telepresence Theory, Measurement & Technology, pp. 115–135. Springer International Publishing, Cham (2015)
10. Mehrabian, A.: Some referents and measures of nonverbal behavior. Behav. Res. Methods Instrum. **1**, 203–207 (1968)
11. Caspi, A., Blau, I.: Social presence in online discussion groups: testing three conceptions and their relations to perceived learning. Soc. Psychol. Educ. **11**, 323–346 (2008)

12. Vareberg, K.R., Westerman, D.: To:-) or to ☺, that is the question: a study of students' initial impressions of instructors' paralinguistic cues. Educ. Inf. Technol. **25**, 4501–4516 (2020)
13. Lee, K.M., Nass, C.: Designing social presence of social actors in human computer interaction. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 289–296. Association for Computing Machinery, Ft. Lauderdale (2003)
14. Yang, J., Sasikumar, P., Bai, H., Barde, A., Sörös, G., Billinghurst, M.: The effects of spatial auditory and visual cues on mixed reality remote collaboration. J. Multimodal User Interfaces **14**(4), 337–352 (2020). https://doi.org/10.1007/s12193-020-00331-1
15. von der Pütten, A.M., Krämer, N.C., Gratch, J., Kang, S.-H.: "It doesn't matter what you are!" Explaining social effects of agents and avatars. Comput. Hum. Behav. **26**, 1641–1650 (2010)
16. Yassien, A., ElAgroudy, P., Makled, E., Abdennadher, S.: A design space for social presence in VR. In: 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, Article 8. Association for Computing Machinery, Tallinn, Estonia (2020)
17. Yorzinski, J.L., Miller, J.: Sclera color enhances gaze perception in humans. PLOS ONE **15**, e0228275 (2020)
18. Kobayashi, H., Kohshima, S.: Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. J. Hum. Evol. **40**, 419–435 (2001)
19. Bailenson, J.N., Beall, A.C., Loomis, J., Blascovich, J., Turk, M.: Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. Hum. Commun. Res. **31**, 511–537 (2005)
20. Fares, M.: Towards multimodal human-like characteristics and expressive visual prosody in virtual agents. In: Proceedings of the 2020 International Conference on Multimodal Interaction, pp. 743–747. ACM, Virtual Event (2020)
21. Pan, X., Gillies, M., Slater, M.: The impact of avatar blushing on the duration of interaction between a real and virtual person. In: Presence 2008: The 11th Annual International Workshop on Presence, pp. 100–106 (2008)
22. Fiore, S., Wiltshire, T., Lobato, E., Jentsch, F., Huang, W., Axelrod, B.: Toward understanding social cues and signals in human–robot interaction: effects of robot gaze and proxemic behavior. Frontiers Psychol. **4** (2013)
23. Bente, G., Rüggenberg, S., Krämer, N.C., Eschenburg, F.: Avatar-mediated networking: increasing social presence and interpersonal trust in net-based collaborations. Hum. Commun. Res. **34**, 287–318 (2008)