Queen Mary University of London

# School of History Assignment Cover Sheet

This coversheet should be completed and form the first page of all assignments.

| | | | | |
|---|---|---|---|---|
| **ITS Username** | ec20165 | | | |
| **Module Code:** | ECS7007P | **Module Title** | Research methods and responsible innovation | |
| **Seminar Teacher** | Dr. George Fazekas | | | |
| **Assignment name** *(e.g. 'Research Essay')* | Coursework 1 – Literature Survey | | **Word Count:** | 3158 |
| **Title:** | Design Thinking: Exploring Non-verbal Communication for Video-mediated Technologies | | | |

## What aspect of this submission would you most like feedback on?

This could be something you are trying particularly to improve, or want to develop for other assignments on this or other modules in the future.

**I have a coversheet from the Disability and Dyslexia Service, and have submitted it alongside this assignment:** ☐

Anonymised copies of student work are occasionally used for teaching. If you do not wish this work to be used in this way, tick the box to the left. ☐

**Assignments must conform to the School of History's submission regulations; those that do not, including late assignments, will be subject to a deduction of marks.**

# Design Thinking: Exploring Non-verbal Communication for Video-mediated Technologies

**Yifan Feng**
for Submission
London, United Kingdom
y.feng@se20.qmul.ac.uk

## ABSTRACT

Video-mediated communication technologies have become a common way of transmitting and transforming human conversation in everyday practice. Its design guidelines root in the processing of information exchange between human and human. The study of non-verbal communication provides a theoretical foundation for examining real-life human interactions in terms of the ways of achieving communicative success without verbal message. Furthermore, it creates opportunities to investigate the subtle patterns behind human-to-human communications (e.g. how people exploit gestural and facial cues) and thereby understand how technology interfaces can alter or be altered between people. Additionally, it helps to reflect upon current effective design principles and devise efficient multimodal human-computer-human dialogue systems. In short, this project explores human-like communication such as how humans employ nonverbal cues to comprehend conversations on a social and cognitive science basis. The ultimate goal of this review is to contribute to making interactive and engaging video-based technologies available to users and highlighting the capability to respond constructively to human social interactions in this design for human-human dialog industry.

### Author Keywords

Video-mediated technologies; human dialog; design thinking; human-human interactions; literature review

### CSS Concepts

**• Human-centered computing~Human computer interaction (HCI)**

## INTRODUCTION

### The promise of co-presence

A fundamental principle to design electronic media for synchronous human interactions is in the efficacy of imitating face-to-face communication [22]. Such a concept relies upon critico-empirical observation of everyday human conversations where social psychologists found two features: social presence [43] and information richness [11] that play a functional role in successfully encoding and decoding messages. In detail, researchers point out that face-to-face dialog enables interlocutors to respond reciprocally with both visual and audible signals [4]. Multi-modal characteristics such as gaze, voice quality, facial expression and other signals enrich possible interactional resources for in-person, onsite communications [22]. Particularly, non-verbal behavioral signals such as hand gestures, forward postures, interpersonal distance and facial expressions contribute to a significant part of successful face-to-face dialog [4, 43]. Experiments such as the map tasks [5] demonstrate the effectiveness of co-presence that can reduce turn-taking per dialog and thereby decrease the possibility of interruptions.

When it comes to video-based communication, mainly there exist two types of issues: anecdotal problems and the misconception of understating human interaction. Date back to 1997, Lewis and Cosier [30] firstly clarified the commercial failure of such technologies and found a pattern that only novices were prone to enjoy its technical affordance. For example, during the Covid-19 pandemic, video-calls are forced to be accepted and widely applied to telemedicine and telehealth [2, 20]. Recent pilot studies revealed that only new customers (i.e. patient) develop (high) satisfaction and frequent users feel the opposite over time due to not only technical issues (e.g. installment of cameras) but a lack of efficient medical consultancy [20, 31]. Regarding current video-supported technologies, designers focus much attention on system parameters such as resolution and frame rate based on a misinterpretation of human talks – the signal processing model of communication [36]. Does more signal data afford more communicative activities?

### Scope of the research

This review supports the idea that a broad range of mechanisms have been introduced into social interactions between humans and the processes are known to be complex [22]. To build telecommunication systems that afford the same variety of interactions, researchers cannot neglect its organization and synthesis procedures [11]. This paper proposes to review nonverbal acts (with or without linguistic meanings) in terms of type, conditions, and its implications. First, this paper explains the meaning of studying human conversations and its importance for video-mediated


*CHI 2020, April 25–30, 2020, Honolulu, HI, USA.*
 ACM ISBN 978-1-4503-6708-0/20/04...$15.00.
DOI: https://doi.org/10.1145/3313831.xxxxxxx

communication technology. Second, the paper elaborates the main approach – systematic literature review on this topic and further present an analysis of the participation of non-verbal cues(acts). The main research questions are: (1) What are nonverbal cues and what are the implications from each kind of nonverbal codes? and (2) What are the lessons for telecom designers could learn from non-verbal interactions?

## DESIGN THINKING: COMMUNICATIVE SUCCESS IN COMMUNICATION TECHNOLOGIES

### The foundation of communication sciences

Proposed in 1948, the mathematical model of human communication ("Shannon-Weaver Data Transmission Model", see Figure 1) was seen as a prominent theory to interpret and analyze human speech [40] and later was applied to probe all human interactions [22]. The formula illustrates that a message is selected by an information source and later converted into signals by a transmitter and sent through a channel. For successful communication, such a message is expected to be decrypted in an identical or close to be identical way for receivers [40]. However, researchers argue that their static and linear model more likely shows the flow of information through a medium rather than an analog to dynamic human communication [1, 4, 22]. Such a model ignores the characteristics of unidirectional and indirect human interactions in the real world of media [1].

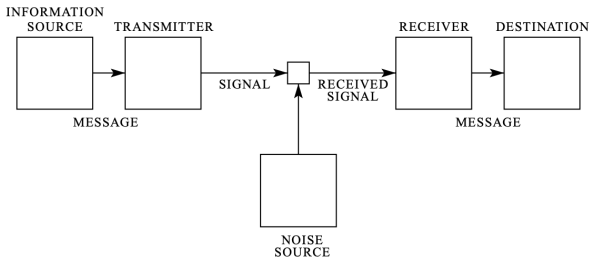


**Figure 1. The schematic diagram of basic communication model [38]**

The weakness of Shannon-Weaver model inspired the focused observation on non-verbal communication in the field of social sciences [14, 17]. By nature, the concept, regardless of strict linguistic meanings, includes all those kinds of human behaviors with referential values and functions (e.g. a set of gestures) [17]. Parallel experimental evidences revealed that these subtle non-verbal messages in certain settings have a close relation to accompanied verbal information that together helps to convey meanings in social communication [4, 34]. As a result, to perform communicative success, one should pay close attention to the functionality of nonverbal aspects in networking and coordinating human activities [35].

### Reflection on emerging technologies

In virtual environments (channeled by video-supported media), the difficulty of employing non-verbal interactions can be seen in the following cases. As aforementioned, face-to-face or co-presence is a vital prerequisite for people producing and perceiving information. Proximity leverages information richness in terms of the type and number of social signals [7]. Researchers [7, 19, 45] noted that chatting via camera and screen makes user feel difficult to: (1) make eye contacts, (2) control body position and gaze, (3) notice peripheral movements, (4) develop side conversations such as whispers, and (5) maintain the joint action ability such as directing an addressee or switching a topic via gaze orientation and finger-pointing. In principle, video-supported communication technologies are said to offer the same or as close as possible experience of co-occurrence meetings and thereby interaction over videos operates in the same way of face-to-face contexts [7, 18]. However, many pilot studies seem to present a difference between face-to-face contacts and video-mediated contacts [7, 9].

To date, the literature regarding the use of video-based technologies in applied settings has become fruitful, much of this focuses on how users can adopt and adapt these technologies from a utility perspective [20, 36, 39]. The reflective study of efficacy and usability of applying new technologies received less attention. This review is in line with Design Thinking Process (see Figure 2) and situated in a deep understanding of non-verbal interactions between human and human. By definition, such an approach integrates human, business and technical factors with a great emphasis on a user-centered viewpoint [6, 15, 38]. In essence, this practice becomes an effective strategy for altering technological changes for potential customers [38].

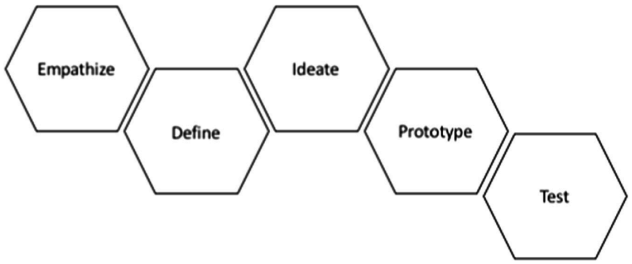


**Figure 2. "Design Thinking Process", initially proposed by Institute of Design at Stanford [15]**

## METHODOLOGY

### SLR and study design

Systematic Literature Review (SLR) refers to a method that stresses a systematic and explicit characteristic of identifying, investigating and evaluating relevant research based on specific research questions, research field or themes [25, 35, 36]. This method not only draws a broader and clearer picture regarding the research topic but also points out the direction for future research [12].

The review is conducted in three stages suggested by Kichenham [27]: (1) planning, (2) execution, and (3) result analysis.

- **Planning:** As said, this review looks into non-verbal cues in real-life human dialogs and its

current design reflection on selected fields such as video-conferencing.

- **Execution:** Literature was gathered through a set of specified search terms in the database of design, computer science and social sciences (see table 1). The initial keywords are determined by the main research topics: nonverbal communication and video-mediated technology and their extensions (not limited) to synonyms such as "nonverbal cues". Besides, as the very first nonverbal communication research started in 1872 by Charles Darwin (see "Nonverbal communication", Wikipedia) publish date is not restricted to/within recent years (i.e. the last 10 years).

| Search terms | Database | Academic disciplines / subject areas |
|---|---|---|
| **Nonverbal cues/acts/behaviors** | Elsevier Science Direct | Computer Science, Psychology |
| **Video-mediated/based technology** | Elsevier Science Direct | Computer science |
| **Nonverbal communication** | Elsevier Science Direct | Social Sciences |
| | ACM Digital Library | Human Factors in Computing Systems |
| **Communication Technology** | ACM Digital Library | Affective computing, computational linguistics |
| | Elsevier Science Direct | Computer Science |

**Table 1. Data collection: Search Terms**

- **Result Analysis:** In total, twenty-seven scholarly projects were selected with a broad time frame (from 1994 to 2020). Mainly, there are three types of publications: book sections (3), peer-reviewed journals (17), and conference papers (7). The chosen articles were analyzed and compared regarding their research objectives, theoretical framework and main findings/discussions. Three categories regarding non-verbal activities are to be presented as theoretical framework in the next section.

## NONVERBAL CUES: FROM THEORY TO PRACTICE

### Human and nonverbal acts

Nonverbal communication is seen as a very general term [17]. Based on Shannon's basic communication model, nonverbal interaction research looks into the particular covariations between a source and a signal [22, 40]. For example, when an observed human behavior (variable one) is altered or shaped by a change in its surrounding environment (variable two), it is concluded that the signal (behavioral change) connotes information about that source (environmental change). To formalize, according to Fernández-Dols [17], the study of nonverbal information is "restricted to the findings of consistent correlations between some observable range of states from the source and some observable states of signals such as bodily postures and movements." Harrigan [21] categorized serval distinctive areas for improving methodological approach to systematically coding and recording such behaviors: (1) facial and vocal behavior, (2) kinesics and gaze behaviors, and (3) proxemics. The following section exemplifies several classic experiments/findings regarding the three categories.

#### *Facial and vocal behavior*

Facial-related research has received the most scholarly attention. The very beginning of facial expressions study was traced back to Darwin's theory of emotions where he proposed that emotions are innate, adaptive and similar across species. His experiments on 60 different emotions demonstrated a universal understanding of discreet expressions such as happiness, sadness, fear and anger, which contributes to the application of human micro expressions (e.g. reading fleeting expressions) in the contemporary psychology and forensic studies. Later on, researchers such as Ekman [16, 21] found that facial expressions can not only relate to emotions but have communicative actions, that is, "referential expressions and conversational signals". His project observed that speakers coordinate eyebrow raises to indicate important elements in their utterance as they use a baton to reinforce vocal stress. Besides, it also showed that such eyebrow signals appear when speaker failed to recall a word (e.g. use filter words 'uh'). On the addressee's side, eyebrows are correspondingly active when information decoding difficulty happens.

Similarly, Duncan et al. [14] researched the correlation between smiling and turn-taking. The project revealed a pattern that the event of addressee's smile occurred at the point where the speaker had finished a complete grammatical unit or a full utterance. Its statistical evidence strongly showed that smile can function as back-channels (e.g. nods, "yeah", and "umm"), which helps to deliver feedback to the speaker without actually speaking out. Chovil [10] in her research of conversational function of facial movements identified the semantic feature of speaker's displays. She believed that such displays illustrated some aspects of the speaker's utterance/dialog at that time (e.g. thinking/remembering face). She also noticed that the addressee's face in the contrast remained less active but more attentive in the dialog. Restricted and simple acts such as

eyebrow raising and lip pressing were recorded in the study and recognized as an additional back channel to smiles.

Nonverbal cues of vocal behavior in speech interactions refer to paralinguistic vocal features other than words, which includes pauses, intonations, loudness and so like [26]. Tempo or the rate of speaking is one the of the popular topics in this field as linguistics found these non-lexical characteristics have the ability to convey informative implicit meanings. For example, McHenry et al. [33] marked a slower uttering rate and lower pitch when the speaker announced bad news as they attempted to integrate the feeling of caring and sympathetic into the event. Another example is filter word such as uh-huh that can alter the order of taking turns. These conversational sounds are taken as indicative signals of emotions, attitudes and affects and help to form and manage interpersonal relationships [26].

*Kinesics and gaze behavior*

Kinesics or body movement research consists of analysis of human actions and positions. Variables in the first category includes head nods, hand movements (e.g. pointing, waving, "ok" sign), shrug and feet (e.g. kicking) [21]. That said, many focuses have been given to gestural analysis and concluded four basic types [4, 8, 34]:

1. Iconic: gestures relate to concrete entity. An example is that an addresser gestures the height of the door when describing that object.
2. Metaphoric: gestures relate to abstract concepts. An example is that an addresser makes heart shape by hands when indicating abstract ideas such as love.
3. Deictic: gestures relate to identification of the location of a referent. An example is that children use index finger as a pointer (the first gesture develops around 10 to 12 months of age).
4. Beats: gestures relate to repetitive movements during the speech and contains no narrative content. An example is that speakers' hands movement follow their speech rhythm.

Researchers found that human gestural behaviors are interactive in nature and therefore cannot be simply understood as a part of speech production [4, 21, 34]. To be clear, experimental observation outlined three features in interactive gestures: (1) particular types of gestures are applied for specified interactivities (2) gestural timing should be considered a sensitive factor during interactivities, and (3) gestural positions are determined by audience structure. Similar to subtle facial and vocal acts, gestures are often deemed expressive of affect, attitude, or intention.

Studies of human positions center on larger units (i.e. the overall bodily movement) that usually compares a beginning position with the change. Such actions are less frequent but easily noticeable. Harrigan [21] proposed several research focuses: (1) overall posture (e.g. sitting), (2) trunk or frontal orientational change (e.g. turning away), (3) trunk lean (e.g. forwards), and (4) arm, leg and foot positions (e.g. uncrossed legs). In particular, the study of head movement received more attention as it is one of the typical and obvious action to code and count during the experiments. Possible acts include nodding, shaking, titling and turning that are often associated with gaze analysis. Essential findings are presented in Harrigan's work [21]:

a. Slight head nod represents a giving and seeking approval or persuasion in both speaking and listening cases.
b. Head shaking was often linked with retrieving negative-valanced words.
c. Rhythmic head movement has association with special features of an utterance (e.g. stress and intonation).

As stated, gaze is deemed as a unique nonverbal cue and usually associated with bodily orientation (particularly head movements). Differently, as a common eye movement, it can channel for both sending and receiving information [5, 21]. Preceding studies [24, 25] on eye-gaze demonstrated that participants can use this technique to guide or regulate turn-taking in conversations and express listeners' emotional and attitudinal feedback to speakers. In detail, some eye movements that appear at the end of a speech serve as synchronization signals to addressees [8]. Moreover, Researchers also found that people can collect information such as other person's intentions via gaze-awareness from our surrounding objects. For example, experimental evidence indicated that infant and young children can follow another person's gaze at an early stage of age (19-20 months) and therefore facilitate the task of language learning [28]. As for adult, gazing acts are an essential indicator in interpersonal communication to show a sign of liking and attraction [18].

*Proxemics*

Proxemics describe the perception, use and framing of space and distance in human interactions [21]. Previous research explores human dialog and proximity in different settings such as business partner meetings, teacher-student collaboration in school, doctor-patient consultancy and recognized some specific conversational space: (1) personal space for discussion between friends, and (2) social space between strangers [19]. The study of proxemics is vital for human communication as human interactions are joint actions [23] and those aforementioned non-speech resources such as body positions/orientations and gestures, facial expressions and eye-gazes are all happening in that shared space. Researchers [8, 13, 43] propose some potential benefits of such research in better understanding human dialogs that it helps to define (1) the role of each participant (i.e. speaker or addressee or over-hearer), (2) their interpersonal or intragroup relationship, (3) level of mutual engagement (i.e. how interest people are in an interactive action), and (4) the starting and ending signals (e.g. timing) of communication.

One of the important findings is F-formation conceptualized by Kendon [24]. This study highlights the use of space as a spatial organization for possible human interactions. Kendon [24] suggested the concept of transactional segment that is the space where people complete actions like talking and looking around. Such space/distance is framed by participants bodily position and orientation and respected by others. In the case of conversation, participants jointly arrange and their transactional segments (or what he called "Overlap space or O-space"), which helps to maintain their communication. Interestingly, such space is dynamic and constantly adjusted by all participants via position changing and so like. These collaborative mutual compensations thereby affect the shape and create some particular patterns for identifying specific relationships: (1) L-arrange for co-operative activities, (2) side-by-side for talking between close friends, and (3) direct face-to-face for greetings.

## DISCUSSION AND CONCLUSIONS

In sum, this review syntheses profound research and major findings in the field of non-verbal communication. Starting with investigation into the strength and weakness of basic communication model (Shannon-Weaver Model), it then suggests how human dialog works in face-to-face settings and the disadvantages of video-mediated communication. The research thinks highly of Design Thinking Process and therefore dives into real-life human interaction where humans are able to exploit subtle nonverbal codes to achieve communicative success. The foundation of this research is based on the idea that successful human interaction is a joint action and requires mutual adjustments [12, 22]. From there, this study looks into three kinds of nonverbal information, namely, facial and vocal behavior, kinesics and gaze behaviors, and proxemics that comes together and assists to send resourceful information for human interactions. Researchers [43] have identified six main functions of amongst aforesaid nonverbal acts: (1) forming impression, (2) managing interaction, (3) expressing emotion, (4) sending relational messages, (5) deceiving and detecting deception, and (6) sending messages of power and persuasion.

Up until now, a subfield of Human-Computer Interaction (HCI) has concentrated on designing for remote presence or activity either for workplace (e.g. video-conferencing) or home (e.g. video-embedded collaborative VR game) and taken into account the value of applying nonverbal cues [19, 32, 41]. That said, evidence suggests that even high-quality video-mediated communication technology failed to fulfill the same benefits as face-to-face, co-present human interaction [13]. To refine current guidelines for devising video-based technologies and afford interactional virtual environments for human users, this research wants to reemphasize the importance of understanding human conversations – which and how subtle cues could appear in collaborative activities in co-present settings and which technical support in video-mediated technologies hinders the transmission of non-verbal messages.


## ACKNOWLEDGEMENTS

The author would like to thank Dr. George Fazekas for his comments and suggestions that significantly enhanced the quality of this research review.


## REFERENCES


[1] Al-Fedaghi, S. 2012. A Conceptual Foundation for the Shannon-Weaver Model of Communication. *International Journal of Soft Computing*. 7, (Jan. 2012), 12–19. https://doi.org/10.3923/ijscomp.2012.12.19.

[2] Anthony, B. 2020. Use of Telemedicine and Virtual Care for Remote Treatment in Response to COVID-19 Pandemic. Journal of Medical Systems. 44, 7 (Jun. 2020), 132. https://doi.org/10.1007/s10916-020-01596-5.

[3] Arminen, I. et al. 2016. Respecifying Mediated Interaction. Research on Language and Social Interaction. 49, 4 (Oct. 2016), 290–309. https://doi.org/10.1080/08351813.2016.1234614.

[4] Bavelas, J. and Gerwing, J. 2007. Conversational Hand Gestures and Facial Displays in Face-to-Face Dialogue. *Social Communication*. Psychology Press. 283–308.

[5] Boyle, E.A. et al. 1994. The Effects of Visibility on Dialogue and Performance in a Cooperative Problem Solving Task. *Language and Speech*. 37, 1 (Jan. 1994), 1–20. https://doi.org/10.1177/002383099403700101.

[6] Buchanan, R. 1992. Wicked Problems in Design Thinking. *Design Issues*. 8, 2 (1992), 5. https://doi.org/10.2307/1511637.

[7] Campbell, J.A. 1998. Participation in videoconferenced meetings: user disposition and meeting context. *Information & Management*. 34, 6 (Dec. 1998), 329–338. https://doi.org/10.1016/S0378-7206(98)00073-1

[8] Cassell, J. et al. 2001. Non-verbal cues for discourse structure. *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics* (USA, Jul. 2001), 114–123.

[9] Chen, C.-H. et al. 2016. Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. *Computers in Human Behavior*. 55, (Feb. 2016), 477–485. https://doi.org/10.1016/j.chb.2015.09.033.

[10] Chovil, N. 1991. Social determinants of facial displays. *Journal of Nonverbal Behavior*. 15, 3 (Sep. 1991), 141–154. https://doi.org/10.1007/BF01672216.

[11] Daft, R.L. and Lengel, R.H. 1986. Organizational Information Requirements, Media Richness and

Structural Design. *Management Science*. 32, 5 (1986), 554–571.

[12] Dikici, A. et al. 2018. Factors influencing the understandability of process models: A systematic literature review. *Information and Software Technology*. 93, (Jan. 2018), 112–129. https://doi.org/10.1016/j.infsof.2017.09.001.

[13] Doherty-Sneddon, G. et al. 1997. Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied*. 3, 2 (1997), 105–125. https://doi.org/10.1037/1076-898X.3.2.105.

[14] Duncan, S. et al. 1979. Strategy signals in face-to-face interaction. *Journal of Personality and Social Psychology*. 37, 2 (1979), 301–313. https://doi.org/10.1037/0022-3514.37.2.301.

[15] Efeoglu, A. et al. 2013. Design Thinking: Characteristics and Promises. *Proceedings 14th International CINet Conference on Business Development and Co-creation*. (2013), 241–256.

[16] Ekman, P. 2009. Darwin's contributions to our understanding of emotional expressions. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 364, 1535 (Dec. 2009), 3449–3451. https://doi.org/10.1098/rstb.2009.0189.

[17] Fernández-Dols, J.-M. 2013. 4 Nonverbal communication: origins, adaptation, and functionality. *Nonverbal Communication*. De Gruyter Mouton. 69–92.

[18] Fullwood, C. 2007. The effect of mediation on impression formation: A comparison of face-to-face and video-mediated conditions. *Applied Ergonomics*. 38, 3 (May 2007), 267–273. https://doi.org/10.1016/j.apergo.2006.06.002.

[19] Grayson, D. and Anderson, A. 2002. Perceptions of proximity in video conferencing. *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2002), 596–597.

[20] Greenhalgh, T. et al. 2020. Video consultations for covid-19. *BMJ*. 368, (Mar. 2020). https://doi.org/10.1136/bmj.m998.

[21] Harrigan, J.A. 2013. 3 Methodology: coding and studying nonverbal behavior. *Nonverbal Communication*. De Gruyter Mouton. 35–68.

[22] Healey, P.G.T. et al. 2018. Editors' Introduction: Miscommunication. *Topics in Cognitive Science*. 10, 2 (2018), 264–278. https://doi.org/10.1111/tops.12340.

[23] Hollan, J. and Stornetta, S. 1992. Beyond being there. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92* (Monterey, California, United States, 1992), 119–125.

[24] Kendon, A. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica*. 26, (Jan. 1967), 22–63. https://doi.org/10.1016/0001-6918(67)90005-4.

[25] Kendon, A. 2010. Spacing and Orientation in Co-present Interaction. *Development of Multimodal Interfaces: Active Listening and Synchrony: Second COST 2102 International Training School, Dublin, Ireland, March 23-27, 2009, Revised Selected Papers*. A. Esposito et al., eds. Springer. 1–15.

[26] Kim, J. et al. 2020. "I Can Feel Your Empathic Voice": Effects of Nonverbal Vocal Cues in Voice User Interface. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, Apr. 2020), 1–8.

[27] Kitchenham, B. 2004. Procedures for Performing Systematic Reviews. *Technical report*. Keele University. 1–26.

[28] Kory Westlund, J.M. et al. 2017. Children use non-verbal cues to learn new words from robots as well as people. *International Journal of Child-Computer Interaction*. 13, (Jul. 2017), 1–9. https://doi.org/10.1016/j.ijcci.2017.04.001.

[29] Kotlyar, I. and Ariely, D. 2013. The effect of nonverbal cues on relationship formation. *Computers in Human Behavior*. 29, 3 (May 2013), 544–551. https://doi.org/10.1016/j.chb.2012.11.020.

[30] Lewis, A.V. and Cosier, G. 1997. Whither video? — pictorial culture and telepresence. *BT Technology Journal*. 15, 4 (Oct. 1997), 64–85. https://doi.org/10.1023/A:1018631512794.

[31] Li, H.-L. et al. 2020. Pilot Study Using Telemedicine Video Consultation for Vascular Patients' Care During the COVID-19 Period. *Annals of Vascular Surgery*. 68, (Oct. 2020), 76–82. https://doi.org/10.1016/j.avsg.2020.06.023.

[32] Maurer, B. et al. 2017. Exploring Gaze and Hand Gestures for Non-Verbal In-Game Communication. *Poster Session*. (2017), 8.

[33] McHenry, M. et al. 2012. Voice analysis during bad news discussion in oncology: reduced pitch, decreased speaking rate, and nonverbal communication of empathy. *Supportive Care in Cancer: Official Journal of the Multinational Association of Supportive Care in Cancer*. 20, 5 (May 2012), 1073–1078. https://doi.org/10.1007/s00520-011-1187-8.

[34] McNeill, D. 2002. Gesture and language dialectic. *Acta Linguistica Hafniensia*. 34, 1 (Jan. 2002), 7–37. https://doi.org/10.1080/03740463.2002.10414607.

[35] Metaxas, D. and Zhang, S. 2013. A review of motion analysis methods for human Nonverbal Communication Computing. *Image and Vision*

*Computing*. 31, 6 (Jun. 2013), 421–433. https://doi.org/10.1016/j.imavis.2013.03.005.

[36] Mlynář, J. et al. 2018. Situated Organization of Video-Mediated Interaction: A Review of Ethnomethodological and Conversation Analytic Studies. *Interacting with Computers*. 30, 2 (Mar. 2018), 73–84. https://doi.org/10.1093/iwc/iwx019.

[37] Moher, D. et al. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ*. 339, (Jul. 2009). https://doi.org/10.1136/bmj.b2535.

[38] Park, H. and McKilligan, S. 2018. A Systematic Literature Review for Human-Computer Interaction and Design Thinking Process Integration. *Design, User Experience, and Usability: Theory and Practice* (Cham, 2018), 725–740.

[39] Patel, R.A. et al. 2013. Visual feedback on nonverbal communication: a design exploration with healthcare professionals. *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare* (Brussels, BEL, May 2013), 105–112.

[40] Shannon, C.E. 1948. A Mathematical Theory of Communication. *Mobile Computing and Communications Review*. 5, (1948), 1–55.

[41] Sundström, P. et al. 2009. Probing the potential of non-verbal group communication. *Proceedings of the ACM 2009 international conference on Supporting group work - GROUP '09* (Sanibel Island, Florida, USA, 2009), 351.

[42] Tiferes, J. et al. 2019. Are gestures worth a thousand words? Verbal and nonverbal communication during robot-assisted surgery. *Applied Ergonomics*. 78, (Jul. 2019), 251–262. https://doi.org/10.1016/j.apergo.2018.02.015.

[43] Vinciarelli, A. et al. 2008. Social signals, their function, and automatic analysis: a survey. *Proceedings of the 10th international conference on Multimodal interfaces* (New York, NY, USA, Oct. 2008), 61–68.

[44] Walther, J.B. 1995. Relational Aspects of Computer-Mediated Communication: Experimental Observations over Time. *Organization Science*. 6, 2 (1995), 186–203.

[45] Whittaker, S. 1995. Rethinking video as a technology for interpersonal communications: theory and design implications. *International Journal of Human-Computer Studies*. 42, 5 (May 1995), 501–529. https://doi.org/10.1006/ijhc.1995.1022.