Computers in Human Behavior 99 (2019) 66–75

Contents lists available at ScienceDirect

Computers in Human Behavior

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

Full length article

Conversational engagement and mobile technology use

Jens F. Binder*, Klaudia Cebula, Sondos Metwally, Michael Vernon, Christopher Atkin, Suvobrata Mitra

Nottingham Trent University, UK

A B S T R A C T

The present work investigated effects of mobility constraints on gesturing and other aspects of conversational engagement when using mobile technology. Based on studies of non-verbal communication and interaction quality, we compared mobile and static forms of video call in structured interviews. Using real-time motion capture and content coding of recordings as well as self-reports, several data sources were considered: hand velocity as an indicator of gesturing, observed levels of engagement and remembered conversational content. Gesturing, in contrast to other indicators of engagement, was reduced under high mobility constraints and was more pronounced under low constraints. Gesturing was further positively related to observed engagement. Additionally, lower constraints resulted in more memories of conversational content. These findings emphasise the relevance of mobility constraints and bodily movement during mediated interaction and call for further model development in computer-mediated communication. Potential for application lies in interactive scenarios such as remote interviewing, testimonials, and relationship maintenance.

1. Introduction

Imagine somebody in a busy inner city district walking at a swift pace, conversing all the while in a raised voice over mobile video call. Imagine the same person in a quiet room sitting at a table conversing over video call on a static computer screen. These examples reflect the simple and observable fact that more and more of our critical conversations, work and personal, are computer-mediated. With the rise of high-bandwidth communication networks and mobile technologies, the use of video calls has become wide-spread. To illustrate these trends, in 2017, the popular video call application Skype passed one billion downloads (The Skype Team, 2017) while Facebook's Messenger recorded 17 billion video call episodes globally (Facebook, 2017).

Global statistics, however, speak little of the actual set-up of devices and environments in which video calls are placed. The focus of the present work is on how conversations and their outcomes differ depending on the physical constraints device use puts on the conversation partners. Whereas the medium, and the technical requirements, are highly similar in the two opening examples, the physical constraints are not, and the level of movement involved is unlikely to be the same. A device that is kept stationary on a table, or a webcam in a static position, is likely to fix the speaker in a seated position. In contrast, a device that is held in hand allows conversations in a wide range of settings and body positions.

Associations between movement, in the sense of bodily activation and physical motion, and the readiness to respond socially can be found frequently in such activities as dancing, hiking, exercising, or performing. At present, however, there is little theory to account for such differences in the field of computer-mediated communication. This may be due to a mis-match between established approaches to computer-mediated communication and the current spread of mobile communication technologies and the resulting spectrum of mobility constraints.

Classical models of mediated communication tend to focus on the editing, sending and decoding of signals, either placing an emphasis on information bandwidth issues (Daft & Lengel, 1986; Short, Williams, & Christie, 1976) or the social-cognitive processes in a reduced-cues environment (Tidwell & Walther, 2002; Walther, 1996). This concerns the longer-term processes necessary for communication partners to establish a common understanding (Spitzberg, 2006; Tidwell & Walther, 2002; Walther, 1996), the inclusion of social validation processes in the absence of reliable online information (Walther, Van Der Heide, Hamel, & Shulman, 2009), the interplay between social presence and interpersonal attraction (Croes, Antheunis, Schouten, & Krahmer, 2016), as well as the intentional and unintentional social signalling via broadcasting technologies (Binder, Howes, & Smart, 2012; Joinson, Houghton, Vasalou, & Marder, 2011).

In contrast, there is a lack of approaches to address the different levels of mobility that are by now available to virtually any technology user world-wide. In the following, we will address this gap by building on research on non-verbal communication (Burgoon, Guerrero, & Floyd, 2010) and embodied communication (Cook & Tanenhaus, 2009). Our core argument is that different levels of mobility imply different constraints on bodily movement which, in turn, have a differential

* Corresponding author. Department of Psychology, Nottingham Trent University, 50 Shakespeare Street, Nottingham, NG1 4FQ, UK.
E-mail address: jens.binder@ntu.ac.uk (J.F. Binder).

https://doi.org/10.1016/j.chb.2019.05.016
Received 9 February 2018; Received in revised form 3 May 2019; Accepted 12 May 2019
Available online 16 May 2019

impact on the overall psychological engagement of conversation partners and interaction outcomes. In pursuing this argument, we develop an integrated theoretical framework as well as a methodological approach suited to accommodate mobility as a novel dimension in CMC. We then go on to report findings from an experiment that compares conversations via video call with different levels of constraints on movement in place.

2. Background: non-verbal communication, movement and interaction quality

Aspects of movement during conversation are clearly related to non-verbal communication, often discussed in the form of facial expression, body posture and gesturing (see Burgoon et al., 2010; Knapp, Hall, & Horgan, 2013). Facial expressions are typically correlated with body gestures (e.g., Castellano, Kessous, & Caridakis, 2008), and as such, both may be affected by mobility constraints in mediated communication. Previous studies have investigated the link between facial expression and interaction quality using detailed content analysis of dyadic interaction. Purvis, Dabbs, and Hopper (1984) found that conversation partners who were skilled at getting others involved and engaged used more attentive facial expressions during interaction. Similarly, Berry and Hansen (2000) reported that independent observer ratings of interaction quality were positively related to visual attention in the form of eye contact and the use of gaze. Other researchers have highlighted the role of facial expression in keeping the conversation positive and ongoing. Kaukomaa, Peräkylä, and Ruusuvuori (2013) found that turn-opening smiles helped to shift the emotional stance of a conversation from neutral to positive and triggered reciprocation. Paggio and Navarretta (2013) concluded that facial signals of different kinds, positive, negative, and attentive, all served as feedback in the pacing of a conversation. While studies on facial expression help to assess the general importance of non-verbal communication for interaction quality, this will become more pronounced when considering body motion as communication next.

Regarding body motion and its relationship with interaction quality, research has focused on upper torso movements, mostly head movement and hand gesturing. Again, the available evidence is based on detailed content analyses of recorded conversations. Hadar, Steiner, and Rose (1985) found that head movement signalled communicative intentions, such that it determined the timing and tempo of positive and negative responses, and that movement helped to synchronize the interactional rhythm. In a similar vein, Paggio and Navarretta (2013) identified a range of head movements as providing instant feedback during interaction. McClave (2000), in some more detail, distinguished between semantic, discourse and communicative functions of head movement. Further, McClave sees the fact that head movement can occur before related verbal content as evidence that movement does not merely reflect the spoken word. Next to head movement, a structuring function has also been ascribed to hand movement, in particular the termination of hand gesticulation which Duncan (1972) found to be a turn-taking signal. Further, as one of the studies most directly concerned with interaction quality, Berry and Hansen (2000) also report on wider body language and found that ratings of interaction quality were positively related to “body openness”, captured as a holistic observer rating.

Next to the functions and correlates of visible body motion, which are about external signalling towards the recipient, some work has looked at the internal functions of movement, i.e., the processes that predominantly concern the speaker. This is of particular relevance in the context of mediated communication and the wide variation in display size and quality. Research has demonstrated that non-verbal activity enables the speaker to articulate their thoughts more effectively, even where these gestures or motions are not visible to the listener (Chawla & Krauss, 1994; Krauss, Chen, & Chawla, 1996). Both facial expressions and gestures are routinely used even when the conversation partner cannot see them to aid in message generation; they act as primers to activate words and concepts that are to be part of the utterance (Burgoon et al., 2010). Gesturing has also been shown to reduce cognitive load while a speaker is generating explanations (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001), which indicates that gesturing absorbs part of the cognitive effort that comes with verbal expressions. In the study by Chawla and Krauss (1994) participants were asked to reproduce recorded statements made by professional actors. They did so with the same overall proportion of gesturing, but also with an increased likelihood of producing gestures classified as lexical movements, those which help to conjure up meanings and definitions. This also means that gesturing can have a direct effect on speech production and on what is being said (see Beilock & Goldin-Meadow, 2010; Chawla & Krauss, 1994). This range of internal and external functions of gesturing identified in previous studies has led to the concept of embodied communication (Cook & Tanenhaus, 2009), a perspective that sees movement as a potentially separate channel for communicating meaning, even across interaction settings. This further underscores the relevance of investigating the possible effects of mobility constraints during conversation.

Regarding the role of technology, our assumption is that communication devices that differ in the level of physical freedom they afford may also differ in the quality or quantity of information they allow to be efficiently transmitted. Some technologies can necessitate and encourage movement, for example in gaming (Graves, Ridgers, & Stratton, 2008; Lindley, Le Couteur, & Berthouze, 2008) and in remote work-settings where gesturing facilitates collaboration (Kirk, Rodden, & Fraser, 2007). Other technology may hinder movement, and thereby social interaction, for example through the use of static displays. A core question, which has not been addressed so far, is whether these varying constraints lead to different levels of movement and thereby to different interaction outcomes. This envisaged role of technology affects both professional and informal settings, corresponding to the everyday ubiquity of CMC, including some situations where comparatively minor variations in outcomes could have substantial consequences. Consider, for example, the remote interviewing for job positions, the delivery of testimonials in legal proceedings, or the upkeep and maintenance of close relationships.

3. The present study: set-up and hypotheses

As discussed in the preceding section, there is convergent evidence for the important and multi-faceted role that movement plays in conversation, coming from studies on face-to-face interaction and more general work on verbal and non-verbal expression. What is lacking, though, is an integration with CMC and a focus on mobile conversation. In the following, we describe such an integrative approach, introducing a methodology and a theoretical stance that is owed substantially to the field of motor cognition (e.g., Jeannerod, 2006) and embodiment (e.g., Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005). Work in these fields focuses on the internal, psychological functions that physical movement, including gesturing, performs in complex cognitive and emotional processes. Studies are often characterised by employing high resolution, real-time motion capture and analysis, which will be triangulated with other data sources here.

Developing a methodology that is suitable for complex interactive settings across varying technologies will enable us to establish the concept of mobility in the study of mediated communication. To this end, this study explores the effects of mobility constraints on different behavioural aspects of conversational interaction: verbal and facial expression as well as gesturing. We are particularly interested in those behavioural aspects that are reflected in interaction quality, and we see the behaviours discussed so far best captured as indicators of conversational engagement, conceptualised here as the psychological involvement and general state of focused activation during the act of conversing (e.g., Bednarik, Eivazi, & Hradis, 2012; Farrant, Maybery, &

Fletcher, 2011). Engagement, put simply, is perceived and experienced as a form of commitment to the conversation. To explore the wider consequences of engagement, this study further investigates the relationship between mobility constraints, indicators of engagement and the memorability of the interaction.

Mobility here is distinguished from movement. The former is a product of the technology and the set-up in which the technology is being used and refers to the general freedom, or affordance, to engage in movement. Movement, then, is understood as the actual physical motion exhibited by technology users. The level of movement displayed during mediated communication is expected to depend on the constraints imposed by the technology. For example, the use of mobile technology may or may not involve walking around, depending on how exactly the device is handled and how the user is positioned.

The type and set-up of communication technologies can be expected to facilitate or inhibit gesturing in two specific ways of interest here. Firstly, technology can prompt users to adopt a particular body posture during conversation depending on device size, display location and so forth. To simplify this aspect, and to allow for a systematic investigation of it, we are concerned with standing and seated postures in the present work, assuming that these two cover the majority of conversations of the average user. A standing posture implies a different pattern of muscular activation (see Jeannerod, 2006), provides more degrees of freedom for overall bodily movement and can lead to further muscular activation in case of walking around. A seated posture, in contrast, restricts most movement to upper torso and arm movements. While this still allows for gesturing, overall activation and facilitation of gesturing should be reduced. Previous research on hand-held device use supports these assumptions. Gustafsson, Johnson, and Hagberg (2010) were able to record higher levels of muscular activity in the upper back when participants were using a mobile phone in a standing as compared to seating posture. Similarly, Straker et al. (2008) found higher levels of upper back muscular activation when children used tablet computers as compared to desktop computers. In addition, they found a higher variability in posture for those using a tablet.

A second way in which technologies can impact on gesturing is through hand engagement. Technology design typically implies hand-held or hands-free use. Hand-held use, again, comes with different patterns of muscular activation (see Gustafsson et al., 2010) which should encourage more movement in the form of gesturing, compared to hands-free use. This means that two combinations of both modes of use provide the strongest contrast: a static mode in which conversation takes place seated and hands-free and a mobile mode in which standing and hand-held use are combined.[1]

Postulating effects of mobility on gesturing raises the question whether direct effects could also be expected on verbal and facial expression. As stated previously, facial expression and gesturing are typically seen to be correlated, and this hold also for verbal and nonverbal conversational behaviour more generally (Burgoon et al., 2010; Castellano et al., 2008). Facial expression and bodily movement are subject to the same initial processing mechanisms in the brain (Meeren, van Heijnsbergen, & de Gelder, 2005), and the same holds for verbal expression and preparedness for movement (Kemmerer & Gonzalez-Castillo, 2010; Raposo, Moss, Stamatakis, & Tyler, 2009). However, research to date has not looked at these associations in the context of technology use. Little is known about how conversation partners respond to constraints imposed by technology. While mobility constraints may affect verbal and facial expression, it is likewise plausible to assume that conversation partners will actively compensate for such obvious loss of information through technology (e.g., by more emphasis, repetition, seeking of clarification). Constraints on movement, in contrast, linked to a less explicit communicative function and directly affected by the set-up of technology as studied here, could be more difficult to overcome. In the absence of further evidence, we postulate that technology effects on movement will be stronger than effects on verbal and facial expression, notwithstanding any correlations among these variables.

A final focus in this work is on the memorability of the conversation. To the extent that movement activates and communicates meanings and concepts (Chawla & Krauss, 1994; Cook & Tanenhaus, 2009), mobility constraints could have longer-lasting effects, beyond the interaction episode. For example, retrieval of autobiographical memories has been shown to be facilitated by movements that match movement connected with the memories (Casasanto & Dijkstra, 2010). Put differently, movement, conceptualised as embodied communication, opens up an additional channel that helps to consolidate content of the interaction. It can therefore be expected that technology-imposed constraints on mobility will have a negative effect on memory for the conversation.

Given the discussion so far, three specific hypotheses were postulated for further investigation. We expected a set-up of standing and hand-held communication to hold fewer constraints on gesturing compared to a set-up of seated and hands-free communication. As a result, the set-up imposing lower constraints should lead to more gesturing during the conversation compared to the one imposing higher constraints (H1). Technology effects on gesturing were further expected to be stronger than any effects on the two other aspects of engagement, verbal and facial expression (H2). Finally, a less constrained set-up was also expected to lead to more memories for conversational content compared to a more constrained set-up (H3).

4. Method

In the present study, participants interacted via video call with a trained interviewer using different set-ups of communication technology. In keeping with the scenarios in the beginning, set-ups were modelled on everyday contexts of communication, as far as possible. Several data sources were combined to capture different aspects of conversational engagement: real-time movement capture, content analyses of video and audio recordings as well as a range of self-report measures.

4.1. Design

A 3 (movement constraints: high vs. medium vs. low) x 3 (conversation topic: self-related information vs. positive past episode vs. negative past episode) mixed-factorial design was employed with movement constraints as a between-subjects and communication topic as a within-subjects factor. The core dependent variables of interest consisted of observed indicators of conversational engagement, i.e., gesturing, facial engagement, verbal engagement, and overall animation. Memory was tested using a simple recall task. In addition, a range of self-report measures were obtained to provide further validation and context for the observed measures. All variables and measures are further described in section 4.4.

4.2. Sample

Data were collected during a three-month period, from May to July 2015. Forty-two participants, 14 male and 28 female, with a mean age of $M = 21.8$ years ($SD = 3.0$) were recruited from departmental participant pools using opportunity sampling. Participants were asked about their past experience of using video calls and estimated usage frequency. Overall, participants had moderate experience with using video calls. From among the sample, 57% reported past use of video calls on a PC, 50% on a smartphone, and 21% on a tablet. Mean self-reported frequencies of use, asked for with reference to the past four weeks, were

[1] A different combination, standing and hands-free use, could be thought of as stimulating rather than inhibiting gesturing. An investigation of all possible constellations was beyond the scope of the present work.

$M_{PC} = 6.67$ ($SD = 11.89$), $M_{PHONE} = 7.38$ ($SD = 7.55$), and $M_{TABLET} = 3.56$ ($SD = 3.05$), for those who confirmed use on these devices. No sample characteristic (age, past use of video calls, usage frequencies, and gender distribution) varied significantly across the between-subjects conditions for movement restriction. All participants were provided with the opportunity to enter in a prize draw at the end of the study for technical equipment of modest value (e.g., memory sticks).

4.3. Set-up and procedure

Standard ethical procedures were adhered to throughout. Participants were randomly assigned to one of three constraints conditions. In the *low constraints condition*, participants were standing in the middle of the room, holding a tablet PC in their non-dominant hand. In the *medium constraints condition*, participants were seated in the same location, holding the tablet PC in their non-dominant hand. A height-adjustable stool was used to discourage a passive seated position. Height was adjusted for each participant such that the upper leg was horizontal in order to avoid prompting counter-balancing torso movements. In the *high constraints condition*, participants were seated in the same manner, but used a laptop PC on a table in front of them. No user input was needed for tablet or laptop. Consequently, the high constraints condition removed any need for movement or muscular activation of the arms. Stool and table were placed at a distance of about 90 cm for all participants in this condition. Display sizes were 8.9 and 11.8 inches for tablet and laptop, respectively. While these set-ups did allow for some uncontrolled variation, e.g., in terms of user-device distance and visual angle, they also mapped directly onto real-life conversational set-ups. Panels a.1, a.2, and a.3 in Fig. 1 illustrate the set-up for low, medium, and high constraints, respectively.

Motion capture was done using two Codamotion CX1 sensor units, positioned at both ends of the laboratory space. The Codamotion system (Rothley, Leicestershire, UK) is an optoelectronic tracker, with infrared emitters placed on the participant and signals picked up by the sensor units. In the present study, emitters were placed on the dominant hand, the hand that was always free across conditions (indicated by arrows in Fig. 1, panel a), and recording was conducted at a sampling rate of 100 Hz. This allowed for continuous and high resolution capture of hand movement for the duration of the experimental procedure. Across all constraints conditions, participants undertook highly similar familiarization tasks. First, they were asked to walk up and down the room. They were then put in position and asked to perform some simple movements to increase and decrease the distance to display in order to establish some general impression of the available movement space. As a final point before actually conversing with another person, they were asked to read out a text that was displayed to them via video call on the device.

Conversations took place in the form of structured interviews. All participants were interviewed remotely by the same female interviewer who was always in the same adjoining room using the same computer equipment. A detailed interview schedule was used for all sessions including rules for follow-up prompts and secondary questions. The schedule had been developed based on the ideas of cognitive interviewing, a technique used to elicit detailed and comprehensive descriptions of memories from interviewees (Memon, Holley, Milne, Koehnken, & Bull, 1994; Memon, Meissner, & Fraser, 2010). Further in line with interviewing techniques, the interviewer had received training prior to data collection on keeping an open and neutral stance. The interview schedule was designed to ensure that all participants would offer longer, continuous answers on three different topics. Conversation topics were pursued in three distinct blocks. In the first interview block, interviewees were led to a point where they provided *self-related information* in the form of short descriptions of themselves and their hobbies. In the second block, memories of a *past positive episode* were elicited by focussing interviewees on "a good time [they had] had during an outdoor activity. Some out-of-doors activity together with a friend, or in a group, or with family. Anything [they felt] positive about." In the third block, memories of a *past negative episode* were elicited by focussing on "some disagreement or debate [they had] had with somebody. Some conversation where [they] had to argue [their] point with a friend, or in a group, or with family. Anything that felt like it could turn into an argument." During each block, the main task of the interviewer consisted in preparing the participant for the topic to come, in stating the topic and in receiving any answers attentively. The interview blocks were marked by short breaks during which participants filled in self-report measures. See Fig. 1, panel c for an overview of blocks and measures. All interviews were recorded using standard recording software for video calls.

Full sessions typically lasted in between 30 and 40 min per participant. Within sessions, mean interview duration, across all three blocks, was $M = 383.0$ s ($SD = 68.1$). Mean durations in seconds for each topic block were $M = 116.5$ ($SD = 32.0$), $M = 112.3$ ($SD = 19.6$) and $M = 154.2$ ($SD = 32.6$) for blocks 1, 2 and 3, respectively.

4.4. Variables and measures

Next to age, gender and video call usage, a range of observational measures were obtained: gesturing from motion capture and the other indicators of conversational engagement, verbal and facial expression as well as general animation, from content coding of interview recordings. In addition, several self-report measures (mood, flow, self-rated expressivity, ratings of the interviewer) were obtained to allow for further investigation of context factors. As a measure of memory performance, a simple recall task was used.

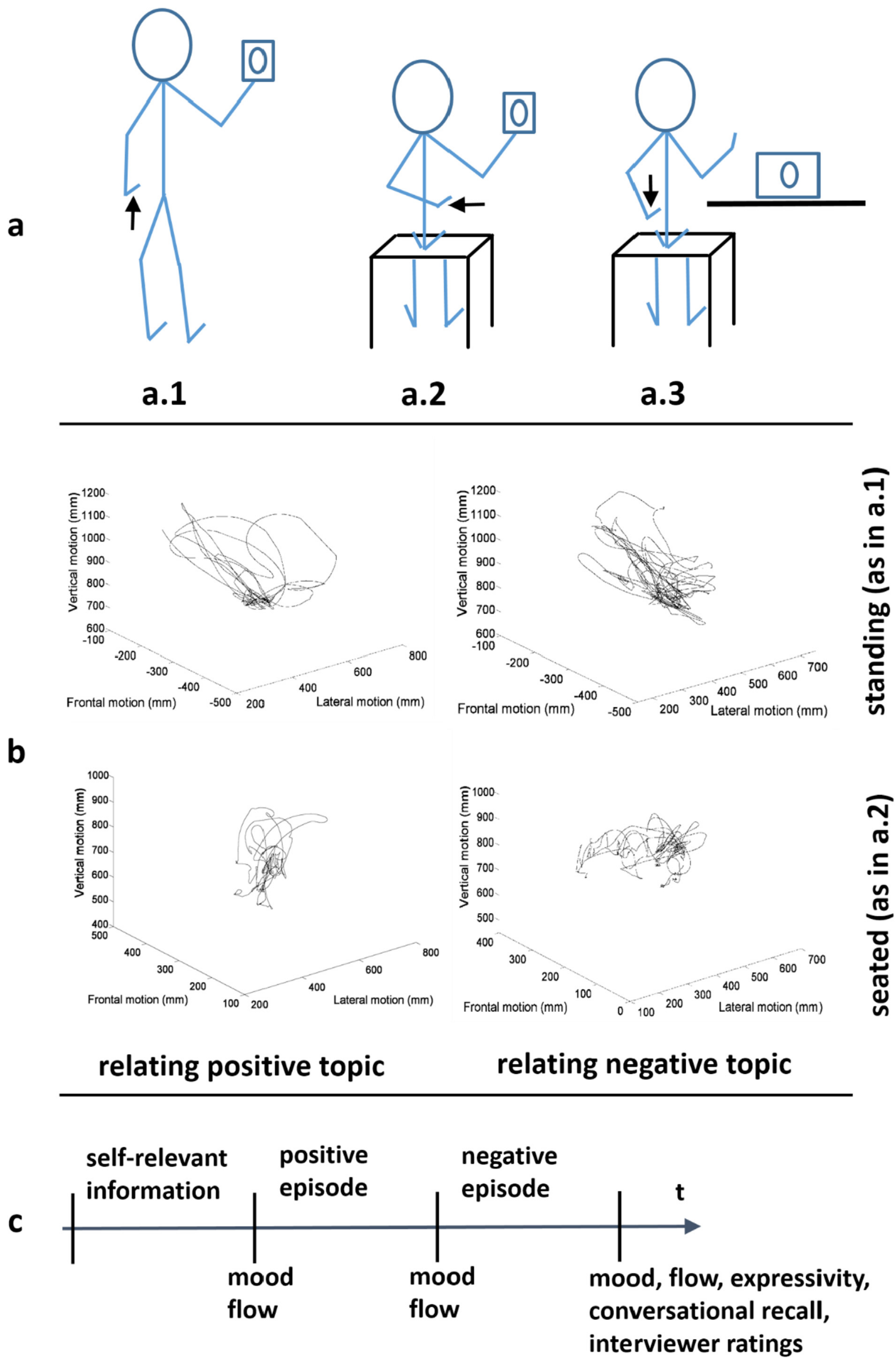
4.4.1. Observational measures

A first inspection of interview recordings and motion capture data showed that participants were most active when delivering responses to the three interview topics. Since the whole interview exercise was designed to elicit these responses, as outlined in the previous section, it was decided to focus on participants' chief responses for the observational measures. Chief responses were defined as the longest, uninterrupted answers provided to the core requests for descriptions for each topic. They typically followed immediately after the main prompt given by the interviewer. Identification of chief responses was facilitated by the fully structured nature of the interviews and the high uniformity of conversation patterns across participants.

A preliminary coding scheme was developed at first, defining variables that mapped onto our conceptualisation of conversational engagement: overall animation, facial engagement, and verbal engagement (gesturing being assessed through motion capture). The aim was to quantify these variables through frequent and regular ratings. A first coder, otherwise unfamiliar with the study, worked with this coding scheme on all interview recordings. During this process, refinements of the variable definitions and pragmatics of coding were added to the scheme.

As a next step in the coding procedure, interrater reliability was investigated by giving half of the recordings (i.e., from 21 participants, seven per constraints condition) to a second rater who had not been involved in the first round of coding. Judgements of both raters showed good levels of agreement across all variables ($r = 0.80$; $\alpha = 0.89$), with levels deemed acceptable for each interview topic: for self-related information $r = 0.68$ ($\alpha = 0.81$), for positive episodes $r = .74$ ($\alpha = 0.84$), for negative episodes $r = .86$ ($\alpha = 0.93$). Based on interrater reliabilities, it was decided at this point in the process to accept the full set of ratings as the data for the main analyses. More information on reliabilities is provided for each variable below. Variables were defined as follows:

Animation. As an indicator of general engagement, animation was defined as the intensity and overall emotionality of facial expression and those nonverbal aspects that indicated a lively, engaged answer.



(caption on next page)

Fig. 1. Panel a depicts the set-up for different levels of movement constraints with arrows indicating emitter position on the dominant hand; panel b provides examples of the hand movement paths obtained to measure gesturing from two participants; panel c shows the general timeline of conversation topic and the self-report measures obtained after each interview block.

Aspects of movement were included in this variable where these were inseparable from facial animation (mostly in cases where gestures supported head posture). Animation was coded on a scale from 1 (passive, inanimate) to 7 (intense, lively), every 10 s during the chief responses that participants gave in each interview. The timeframe was chosen to ensure that even shorter responses were adequately represented by several measures. Overall interrater reliability for this variable was $r = 0.79$ ($\alpha = 0.88$).

Facial Valence. As an indicator of facial engagement, valence was defined as a supplementary aspect to animation, focussing solely on positivity of facial expression. Given a conversation setting that was purposefully kept in a positive tone, instances of discernible negative facial valence were very rare. The scale was anchored accordingly such that valence was coded on a scale from 1 (neutral) to 7 (very positive), every 10 s during the chief responses that participants gave in each interview block. Overall interrater reliability for this variable was $r = 0.77$ ($\alpha = 0.87$).

Dynamic Description. As an indicator of verbal engagement, dynamic description referred to several aspects of action in a description: behaviours being carried out, the unfolding of a scene, dialogue, movement or other sequences of events. Dynamic description was coded on a scale from 1 (very little going on, no action markers) to 7 (lots going on, full of action markers). Given that complex propositions can take up significantly less than 10 s (Kitsch & Keenan, 1973), it was decided to obtain ratings every 5 s during the chief responses that participants gave in each interview block. Overall interrater reliability for this variable was $r = 0.65$ ($\alpha = 0.79$).

Gesturing. As an indicator of gesturing, hand velocity was computed from path lengths and talking time for the chief responses that the participants gave in each interview block. The movement signal was obtained from the back of the dominant hand, which was always free to use across all conditions. Velocity was scaled as mm/s.

4.4.2. Self-report measures

Next to memory performance, two self-report measures on mood and flow were included to check whether the manipulations would have any unintentional effects on emotional and motivational states. Further, perceptions of the interviewer were assessed to check whether this aspect of the setting would show any undesirable variability across conditions. Finally, a measure of expressivity was designed to investigate whether participants would show any awareness of externally observed levels of engagement.

Conversational Recall. As a simple memory task, interviewees were instructed to think back to the conversation and to write down all the things they could remember from the situation, using keywords rather than full text. Space was provided on the page for up to ten entries, in the form of ten separate lines. Recall was then measured as the number of entries made ($M = 6.23$; $SD = 2.10$). Entries were checked by two independent coders for duplications, showing high levels of agreement ($r = 0.92$; $\alpha = 0.97$). The few cases of discrepancies (two) were resolved by a third party.

Mood. A scale composed of four positive and four negative mood adjectives designed to assess momentary affect, taken from Parkinson, Briner, Reynolds, and Totterdell (1995), was used to measure mood ("cheerful", "involved", "alert", "calm", "tired", "tense", "depressed", "disinterested"). Adjectives were rated with options ranging from 1 (I felt like this not at all) to 7 (I felt like this very much), and were recoded and averaged to form indices with higher values indicating more positive mood. Reliabilities across the three measurement points were mixed: $\alpha_{t1} = 0.62$, $\alpha_{t2} = 0.66$, and $\alpha_{t3} = 0.79$. Since all attributes showed positive item-total correlations, and in the absence of further indicators for item reduction, the measure was used without further modification.

Flow. Eight items representing two core factors of flow states, autotelic experience (e.g., "The conversation left me feeling great.") and concentration (e.g., "My attention was focused entirely on what I was doing."), were taken from the Flow State Scale (Jackson & Marsh, 1996; items included were numbers 5, 9, 14, 18, 23, 27, 32, 36) to capture those aspects of flow that seemed most applicable to the comparatively brief communication experience. Items were rated for agreement with options ranging from 1 (not at all) to 7 (very much), and were recoded and averaged to form reliable indices ($\alpha_{t1} = 0.69$, $\alpha_{t2} = 0.77$, $\alpha_{t3} = 0.84$) with higher values indicating higher levels of flow experience.

Ratings of the Interviewer. The interviewer was rated on a number of attributes ("trustworthy", "likeable", "attractive" as well as four attributes representing warmth in Fiske, Cuddy, Glick, and Xu (2002): "warm", "good natured", "sincere", "tolerant"), with options ranging from 1 (not at all) to 7 (very much). For warmth ($\alpha = 0.78$) items were averaged to form an index with higher values indicating higher levels of warmth.

Expressivity. Six researcher-generated items intended to capture interviewees own perceived ability to make themselves understood during the conversation: "I was able to get my points across.", "I felt what I said was inadequate.", "I left an accurate impression of myself.", "I was not able to describe things properly.", "I was able to express myself as I wanted.", and "I was struggling with delivering my message.". Items were rated for agreement with options ranging from 1 (not at all) to 7 (very much), and were recoded and averaged to form a reliable index ($\alpha = 0.75$) with higher values indicating more perceived expressivity.

4.5. Data analysis

In order to take account of the fully crossed two-factorial design, mixed-factorial analysis of variance (ANOVA) was used for the core analyses. When background variables were included as statistical controls, analysis of covariance (ANCOVA) was employed, to partial out overlap between the dependent variable and control variables. Bonferroni tests were used for post-hoc comparison of the three experimental conditions, to avoid error inflation. Multivariate analysis of variance (MANOVA) was used in one instance where conceptually interrelated dependent variables, all of the ratings associated with the interviewer, were scrutinised for any statistical effects in an exploratory fashion.

5. Results

5.1. Preliminary analyses

A general selection criterion was applied to participants imposing lower and upper limits on gesturing of 10 mm/s and 1000 mm/s, respectively. In addition, due to signal loss in cases of emitter occlusion, complete sets of velocity measures, for all three topics, could only be obtained for 28 participants. In seven cases where only one measure was missing, it was therefore decided to use mean imputation. This meant that the main analyses could be based on 35 participants and 105 data points for gesturing.[2] Mean scores for all dependent measures for

[2] Mean imputation was based on gesturing measurements within the same topic, but across all three conditions. Imputation therefore supported a conservative estimate of between-condition effects.

Table 1
Mean scores and standard deviations for all variables obtained for each of the conversation topics.

DV	Set-Up and Conversation Topic								
	Standing, device in hand			Seated, device in hand			Seated, hands free		
	Self-related	Positive outdoors	Disagreement	Self-related	Positive outdoors	Disagreement	Self-related	Positive outdoors	Disagreement
Gesturing	292.48 (274.70)	338.72 (298.06)	349.98 (203.91)	139.82 (195.53)	170.14 (140.39)	237.37 (189.25)	126.80 (107.28)	99.86 (119.33)	300.21 (223.60)
Facial Expression	4.46 (0.60)	4.58 (0.97)	4.44 (0.82)	4.59 (0.40)	4.61 (0.49)	4.31 (0.84)	4.27 (0.53)	4.23 (0.60)	3.86 (0.65)
Dynamic Description	4.63 (0.68)	4.78 (0.68)	5.08 (0.70)	4.57 (0.65)	4.70 (0.59)	5.14 (0.35)	4.60 (0.55)	4.80 (0.57)	5.09 (0.27)
Animation	4.56 (0.39)	4.75 (0.45)	4.38 (0.62)	4.36 (0.94)	4.43 (1.09)	4.47 (0.55)	4.31 (0.63)	4.38 (0.65)	4.01 (0.72)
Mood	5.64 (0.77)	5.81 (0.79)	5.06 (0.92)	5.95 (0.60)	6.10 (0.71)	5.80 (0.75)	5.60 (0.65)	6.04 (0.52)	5.03 (0.88)
Flow	4.54 (0.71)	4.97 (0.80)	4.17 (0.96)	5.06 (0.80)	5.19 (0.87)	4.71 (1.01)	4.88 (0.71)	5.47 (0.73)	4.46 (1.12)

Note. Standard deviations in parentheses.

Table 2
Mean scores and standard deviations for all variables assessed after the interview.

DV	Set-Up		
	Standing, device in hand	Seated, device in hand	Seated, hands free
Conversational Recall	7.31 (2.32)	5.00 (1.24)	6.08 (1.85)
Expressivity	4.33 (1.08)	4.80 (1.08)	4.68 (0.83)
Likable	6.08 (0.86)	6.30 (0.68)	6.23 (0.73)
Attractive	6.08 (0.64)	5.80 (1.23)	6.31 (0.75)
Trustworthy	5.38 (1.39)	5.40 (1.17)	6.00 (0.82)
Warmth	6.10 (0.63)	6.23 (0.69)	6.21 (0.75)

Note. Standard deviations in parentheses.

the final sample are summarised in Tables 1 and 2.

It was then investigated whether the interviewer had left different impressions at different levels of mobility constraints. All ratings of the interviewer were therefore subjected to a MANOVA. This yielded no significant effects for the constraints factor, neither overall nor for a specific dependent measure, and it was concluded that interviewer effects could have played a minor role at best.

Further, affective responses to mobility constraints were explored. Mood and flow ratings were subjected to 3 × 3 ANOVAs with constraints as between-subjects and conversation topic as within-subjects factor. The only significant effects found were for topic. The most positive mood ratings were observed for the relating of positive episodes ($M = 6.02$, $SD = 0.68$), followed by self-related information ($M = 5.77$, $SD = 0.66$) and negative episodes ($M = 5.30$, $SD = 0.90$): $F(2, 62) = 18.29$, $p < .001$, $\eta^2 = 0.37$. Similarly, participants reported highest levels of flow experience when relating positive episodes ($M = 5.28$, $SD = 0.76$) compared to self-related information ($M = 4.86$, $SD = 0.74$) and negative episodes ($M = 4.49$, $SD = 0.99$): $F(2, 62) = 12.98$, $p < .001$, $\eta^2 = 0.30$. It was concluded that, although mood and flow were sensitive to the valence of the topic of conversation, affective responses to mobility constraints did not play a discernible role here.

5.2. Mobility constraints on conversational engagement

Participant gender and video call usage were correlated with all indicators of engagement (gesturing, animation, facial valence, dynamic description). Only two significant correlations emerged, between call usage and gesturing relating a negative episode ($r = 0.41$, $p = .02$) as well as between call usage and facial valence relating a positive episode ($r = 0.34$, $p = .04$). In addition, negative correlations between gender and gesturing relating self-relevant information ($r = -0.33$, $p < .06$) as well as animation relating self-relevant information ($r = -0.33$, $p < .06$) indicated lower scores for female participants compared to males in the first interview block for gesturing and animation. Gender and call usage were therefore entered as covariates in all analyses of engagement measures.

Regarding the effects of mobility constraints, H1 stated that set-ups imposing lower constraints would lead to more gesturing compared to those imposing higher constraints. A 3 × 3 mixed-factorial ANCOVA was conducted on gesturing with topic as within-subjects and level of constraints as between-subjects factor, as well as gender and call usage as covariates. Results indicated a significant main effect for constraints ($F(2, 30) = 5.56$, $p = .01$, $\eta^2 = 0.27$). Gender and topic showed a significant interaction ($F(2, 60) = 3.42$, $p = .04$, $\eta^2 = 0.10$). Separate t-tests showed that levels of gesturing were somewhat higher for male participants compared to females for the first topic ($t(33) = 2.03$, $p = .05$), but not for the other topics ($ps > .02$). In addition, call usage in the ANCOVA accounted for a significant proportion of variance in gesturing ($F(1, 30) = 7.19$, $p = .01$, $\eta^2 = 0.19$). No other significant effects were found. In line with H1, highest estimated mean scores were obtained in the low constraints condition ($M = 345.70$, $SD = 43.15$), followed by the medium ($M = 189.61$, $SD = 47.25$) and high ($M = 143.44$, $SD = 47.44$) constraints conditions.[3] Post-hoc Bonferroni tests on estimated means indicated a significant difference between high and low constraints ($p = .01$) as well as a marginally significant difference between high and medium constraints ($p = .06$). In sum, this means that free-standing, hand-held conversations led to highest levels of gesturing, thus supporting H1. Similar 3 × 3 ANCOVAs were carried out for the two components of gesturing, the duration of interviewee replies and path lengths. No similar effects for mobility constraints were found ($ps > .45$).

Regarding the effects of mobility constraints on other aspects of conversational engagement, H2 stated that such effects would be smaller in comparison to gesturing. A series of 3 × 3 mixed-factorial ANCOVAs was conducted with gender and call usage as covariates, topic as within-subjects and level of constraints as between-subjects factor on the remaining measures of engagement: animation, facial valence and dynamic description. No significant effects for constraints were found in these analyses.

For animation, the only significant effects were for the covariates. Gender ($F(1, 30) = 6.54$, $p = .2$) and call usage ($F(1, 30) = 6.32$, $p = .02$) both accounted for significant portions of variance in animation. For facial valence, a significant main effect for topic ($F(2, 60) = 4.00$, $p = .02$, $\eta^2 = 0.12$) was found, and Bonferroni tests showed that negative episodes elicited significantly less positive facial valence ($M = 3.84$, SD = 0.55) than self-related information ($M = 4.44$, SD = 0.53) or positive episodes ($M = 4.21$, SD = 0.77, all $ps < .01$). In addition, topic and call usage interacted ($F(2, 60) = 3.36$, $p = .04$,

[3] The main effect for constraints remained stable when using a smaller sample without mean imputation: $F(2, 23) = 5.35$, $p = .01$.

$\eta^2 = 0.10$), in line with the positive correlation reported earlier. No other significant effects were found for facial valence. For dynamic description, a significant main effect for topic ($F(2, 52) = 3.80$, $p = .03$, $\eta^2 = 0.13$) was found, with Bonferroni tests showing less dynamic description for self-related information ($M = 4.61$, SD = 0.64) compared to positive episodes ($M = 5.10$, SD = 0.50) and negative episodes ($M = 5.07$, SD = 0.52, all $ps < .01$). No other significant effects were found for dynamic description.[4] Overall, the pattern of constraint effects, significant for gesturing, non-significant for animation, facial valence and dynamic description, lends support to H2.

Further associations among aspects of engagements were investigated through bivariate correlations between gesturing, animation, facial valence, dynamic description, and self-reported expressivity. Gesturing was significantly related to the more general aspect of animation when relating self-relevant information ($r = 0.42$, $p = .01$) and positive episodes ($r = 0.39$, $p = .02$). Higher levels of gesturing were mirrored in higher levels of observed animation. Additionally, gesturing was related to facial valence when positive episodes were related ($r = 0.35$, $p = .04$). Higher levels of gesturing were accompanied by more positive facial expressions. Finally, gesturing was related to self-rated expressivity, again for positive episodes ($r = -0.59$, $p < .001$). The more participants engaged in gesturing during this conversation topic, the less they felt able to make themselves understood. In sum, although other indicators of conversational engagement were in part sensitive to the topic of conversation and were correlated with gesturing, they were unaffected by mobility constraints. In addition, gesturing seemed to be reflected in more, rather than less, effort to express meaning accurately.

5.3. Mobility constraints on conversational recall

Regarding the effects of mobility constraints on memory of the conversation, H3 stated that lower mobility constraints would lead to more memories compared to higher constraints. A 1 × 3 ANOVA was conducted comparing high, medium and low levels of constraints for conversational recall. The effect was significant ($F(2, 31) = 3.96$, $p = .03$, $\eta^2 = 0.20$), and Bonferroni post-hoc tests indicated that low ($M = 7.31$, $SD = 2.32$) mobility constraints led to significantly more recall than medium constraints ($M = 5.00$, $SD = 1.25$), although high mobility constraints ($M = 6.18$, $SD = 1.99$) did not differ significantly from either of the other two conditions.[5] Mean scores are displayed in Fig. 2. Thus, H3 was partially supported.

To pursue the question whether the effects of mobility constraints on recall could be accounted for by gesturing, an additional 1 × 3 ANCOVA was carried out with gesturing during all three topics as covariates. While this did reduce the main effect to non-significance ($F(2, 28) = 2.62$, $p = .09$, $\eta^2 = 0.16$), the changes could not be attributed to variance accounted for by the covariates. The pooled influence of the three gesturing measures remained non-significant ($F(3, 28) = 1.06$, *ns*). As such, no evidence could be obtained that gesturing mediated the effects of mobility constraints on recall.

6. Discussion

The present work outlines a perspective that integrates body movement, mobile technology and conversational engagement with the aim of addressing psychological effects of mobile-mediated interaction. A method and procedure are presented that allow for accurate measurement of movement in social interactive settings with high ecological validity. Affordance of movement is a dimension hitherto mostly ignored in CMC, but one that provides a departure from scrutinising communication channels for information bandwidth. In fact, one and the same channel can come at different levels of mobility, which has wider psychological implications, as our findings suggest.

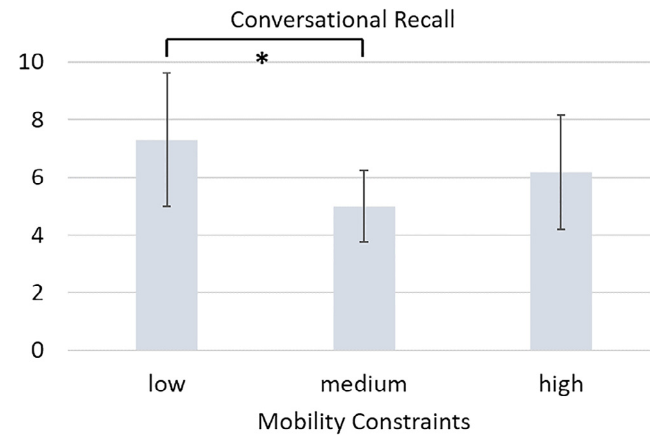


Fig. 2. Number of items remembered from the conversation at different levels of mobility. Error bars represent standard deviations. Conditions that differ significantly from each other are indicated by asterisk ($p < .05$).

Evidence was found for the main hypotheses, H1 and H2, relating to gesturing and mobility constraints. Gesturing, as expected, was greater when under lower constraints, at least when a standing posture with device in hand is compared to a seated posture with hands free. While a seated posture with device in hand led to moderate levels of gesturing, as hypothesised, these were not statistically significant from those in the other conditions.

At the same time, only gesturing was sensitive to the manipulation. Other indicators of engagement, general animation, facial expression and verbal expression, remained unaffected. Further, it is unlikely that effects on gesturing are attributable to affective-motivational factors. Reported mood and flow levels were unaffected by the set-up of technology. Importantly, the set-up of technology did not by necessity require more or less hand movement. Measures were obtained from the arm that was not involved in the handling of the communication technology. Further, our measure of gesturing was a velocity measure, chosen as an indicator of activation and alertness, not a measure of the overall amount of movement. We therefore see our findings as indicating that mobile technology can bring about direct changes in conversation partners' physical engagement.

This is not to mean that gesturing was completely isolated from other indicators of engagement. Positive associations were found between gesturing and general animation as well as positive facial expression. With higher levels of gesturing participants also tended to report lower levels of expressivity. In other words, gesturing increased when interviewees found it difficult to get their points across and make themselves understood. This supports the view that gesturing supports more challenging explanations – not necessarily for the listener (next to no gesturing was discernible in the video call recordings in our study), but for cognitive elaboration on the speaker's side (Chawla & Krauss, 1994; Goldin-Meadow et al., 2001; Krauss et al., 1996). The study did not include an observational measure of people's actual expressive difficulties when talking. It is unlikely, though, that the manipulation of mobility had an effect on people's ability to relate information. Verbal expression, although variable across conversation topics, was not sensitive to mobility constraints.

The present work is the first to relate such constraints to conversational engagement and interaction outcomes under conditions of immediate ecological validity. This inevitably means that the manipulation of mobility constraints used here was not without shortcomings. A range of factors of cognitive relevance, such as visual angle, actual

[4] Removing covariates and conducting corresponding ANOVAs did not alter the picture for constraints effects.

[5] Adding gender and call usage as covariates did not change the picture for conversational recall. Covariates were non-significant and the main effect remained virtually unchanged.

volume of speech, physical differences between tablet and laptop (e.g., keyboard) could not be controlled for. While no immediate signs of experimental artefacts emerged during data analysis, these issues should be addressed in future studies on technology and communicative engagement. This way, it may also be possible to assess more fine-grained differences among set-ups and separate out the effects of different levels of mobility constraints.

The choice of set-up is related to the choice of participants. While the sample used here did allow for a comparison of experimental conditions, it is possible that the situation did not suit all participants equally, in particular when it comes to experience with using video calls. It may have been a more desirable option to use highly experienced participants only, which could help to avoid any feelings of awkwardness in handling the technology. This, however, would have partly run counter to our aim of creating "normal situations", with participants who are neither highly trained nor full novices. Sample characteristics in the present study showed moderate levels of video call experience. In addition, most parts of the structured interview were intended to lead interviewees securely up to the point where they would deliver some detailed descriptions. It should also be emphasised that the requirements for motion capture, carrying sensors and moving within the defined space that was covered by detectors, are in themselves unusual, and that preliminary exercises were used to mitigate these factors. Still, for future studies some pre-selection of participants according to their communication experiences is to be recommended.

Mobility constraints further showed an effect on conversational recall. Given that memory is rarely considered in studies on social interaction, this general finding in itself opens up novel avenues for further investigation. In line with H3, low constraints, standing with tablet in hand, led to the highest number of items remembered from the conversation, and this was significantly different from medium constraints, when participants were seated with tablet in hand. Contrary to expectations, however, participants who conversed seated via a laptop did not show significantly different recall from the tablet conditions. Based on previous research linking motor activation to memory (Casasanto & Dijkstra, 2010; Chawla & Krauss, 1994; Cook & Tanenhaus, 2009), increased motoric engagement under lower mobility constraints was thought to increase elaboration and encoding of conversational content. This theoretical argument, however, could not be supported from our data since gesturing did not account for mobility effects on recall when used as a covariate.

At this point, only speculations can be offered to account for the mixed findings relating to H3. It is possible that a seated posture with hands free allowed for additional activation and elaboration that benefited memory. For example, in this condition, there is no need to actually handle a device and to manage visual contact with the interviewer. In other words, cognitive interference may have been lower with hands freed up. Our recommendation for future research would therefore be to include measures of cognitive load and attention, preferably together with additional measures of movement and bodily activation.

The behavioural aspects of conversation as studied here are all related to commitment, attention, and the regulation of affect as well as content. We have therefore used the general term conversational engagement to refer to states of focused activation during conversation (see also Bednarik et al., 2012; Farrant et al., 2011). According to our findings, mobile technology can have a direct effect on engagement via gesturing, while indirect effects on other indicators of engagement are likely to be the case. Importantly, the effects seem to be based on the internal functions of gesturing, those relevant to the speaker, rather than external signalling. Our findings therefore suggest that engagement could be affected in an immediate and fundamental way that is not easily noticed and compensated for by technology users.

7. Conclusion

Understanding movement is increasingly important as CMC routinely takes place in situations that come with different user constraints on mobility. The effects of such constraints are likely to be of wider significance in all social interactions that involve impression formation and the establishment of common ground, trust and credibility. In the context of mobile interaction, a range of applied settings can be identified for further work, from remote interviewing for job positions, to legal contexts where witness and testimonials are to be obtained, to the upkeep and maintenance of close relationships.

Aside from applications, the concept of conversational engagement that we have proposed and operationalised here presents an empirically grounded theoretical device that can be used to expand the scope of CMC models, and relate CMC to wider theories of verbal and non-verbal communication. As mobile technology becomes increasingly common as a CMC platform, and ever more critical human interactions are conducted on it, linking users' conversational engagement to their movement and mobility constraints becomes an essential aspect of a theoretical framework for how people present themselves and how they evaluate others' presentation. It is worth noting that the presented framework for conversational engagement includes measures of conversational effectiveness, and this aspect may be further developed as mobile-platform CMC becomes an increasingly important and ubiquitous mode of embodied communication (Cook & Tanenhaus, 2009).

References

Bednarik, R., Eivazi, S., & Hradis, M. (2012). Gaze and conversational engagement in multiparty video conversation: An annotation scheme and classification of high and low levels of engagement. *Proceedings of the 4th workshop on eye gaze in intelligent human machine interaction* (pp. 10–16). . ACM https://doi.org/10.1145/2401836.2401846.

Beilock, S. L., & Goldin-Meadow, S. (2010). Gesture changes thought by grounding it in action. *Psychological Science, 21*, 1605–1610.

Berry, D. S., & Hansen, J. S. (2000). Personality, nonverbal behavior, and interaction quality in female dyads. *Personality and Social Psychology Bulletin, 26*, 278–292.

Binder, J. F., Howes, A., & Smart, D. (2012). Harmony and tension on social network sites. *Information, Communication & Society, 15*, 1279–1297. http://doi.org/10.1080/1369118X.2011.648949.

Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2010). *Nonverbal communication.* Boston, MA: Allyn & Bacon.

Casasanto, D., & Dijkstra, K. (2010). Motor action and emotional memory. *Cognition, 115*, 179–185. http://doi.org/10.1016/j.cognition.2009.11.002.

Castellano, G., Kessous, L., & Caridakis, G. (2008). Emotion recognition through multiple modalities: Face, body gesture, speech. *Affect and emotion in human-computer interaction* (pp. 92–103). Berlin: Springer.

Chawla, P., & Krauss, R. M. (1994). Gesture and speech in spontaneous and rehearsed narratives. *Journal of Experimental Social Psychology, 30*, 580–601.

Cook, S. W., & Tanenhaus, M. K. (2009). Embodied communication: Speakers' gestures affect listeners' actions. *Cognition, 113*, 98–104.

Croes, E. A., Antheunis, M. L., Schouten, A. P., & Krahmer, E. J. (2016). Teasing apart the effect of visibility and physical co-presence to examine the effect of CMC on interpersonal attraction. *Computers in Human Behavior, 55*, 468–476. http://doi.org/10.1016/j.chb.2015.09.037.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*, 554–571. http://doi.org/10.1287/mnsc.32.5.554.

Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology, 23*, 283–292.

Facebook (2017). *Messenger's 2017 Year in review.* Available at https://newsroom.fb.com/news/2017/12/messengers-2017-year-in-review/, Accessed date: 31 January 2019.

Farrant, B. M., Maybery, M. T., & Fletcher, J. (2011). Socio-emotional engagement, joint attention, imitation, and conversation skill: Analysis in typical development and specific language impairment. *First Language, 31*, 23–46. http://doi.org/10.1177/0142723710365431.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902. https://doi.org/10.1037/0022-3514.82.6.878.

Goldin-Meadow, S., Nusbaum, H., Kelly, S., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12*, 516–522. http://doi.org/10.1111/1467-9280.00395.

Graves, L. E., Ridgers, N. D., & Stratton, G. (2008). The contribution of upper limb and total body movement to adolescents' energy expenditure whilst playing Nintendo Wii. *European Journal of Applied Physiology, 104*, 617. http://doi.org/10.1007/s00421-008-0813-8.

Gustafsson, E., Johnson, P. W., & Hagberg, M. (2010). Thumb postures and physical loads during mobile phone use–A comparison of young adults with and without musculoskeletal symptoms. *Journal of Electromyography and Kinesiology, 20*, 127–135. http://doi.org/10.1016/j.jelekin.2008.11.010.

Hadar, U., Steiner, T. J., & Rose, F. C. (1985). Head movement during listening turns in conversation. *Journal of Nonverbal Behavior, 9*, 214–228.

Jackson, S. A., & Marsh, H. W. (1996). Development and validation of a scale to measure optimal experience: The Flow State Scale. *Journal of Sport & Exercise Psychology, 18*, 17–35. http://doi.org/10.1123/jsep.18.1.17.

Jeannerod, M. (2006). *Motor cognition: What actions tell the self.* Oxford: Oxford University Press.

Joinson, A. N., Houghton, D. J., Vasalou, A., & Marder, B. L. (2011). Digital crowding: Privacy, self-disclosure, and technology. *Privacy online* (pp. 33–45). Berlin: Springer. http://doi.org/10.1007/978-3-642-21521-6_4.

Kaukomaa, T., Peräkylä, A., & Ruusuvuori, J. (2013). Turn-opening smiles: Facial expression constructing emotional transition in conversation. *Journal of Pragmatics, 55*, 21–42.

Kemmerer, D., & Gonzalez-Castillo, J. (2010). The two-level theory of verb meaning: An approach to integrating the semantics of action with the mirror neuron system. *Brain and Language, 112*, 54–76. http://doi.org/10.1016/j.bandl.2008.09.010.

Kirk, D., Rodden, T., & Fraser, D. S. (2007). Turn it this way: Grounding collaborative action with remote gestures. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1039–1048). . ACM http://doi.org/10.1145/1240624.1240782.

Knapp, M., Hall, J., & Horgan, T. (2013). *Nonverbal communication in human interaction.* Boston, MA: Cengage Learning.

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? *Advances in Experimental Social Psychology, 28*, 389–450. http://doi.org/10.1016/S0065-2601(08)60241-5.

Lindley, S. E., Le Couteur, J., & Berthouze, N. L. (2008). Stirring up experience through movement in game play: Effects on engagement and social behaviour. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 511–514). . ACM http://doi.org/10.1145/1357054.1357136.

McClave, E. Z. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics, 32*, 855–878.

Meeren, H. K., van Heijnsbergen, C. C., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, 102*, 16518–16523. http://doi.org/10.1073/pnas.0507650102.

Memon, A., Holley, A., Milne, R., Koehnken, G., & Bull, R. (1994). Towards understanding the effects of interviewer training in evaluating the cognitive interview. *Applied Cognitive Psychology, 8*, 641–659. https://doi.org/10.1002/acp.2350080704.

Memon, A., Meissner, C. A., & Fraser, J. (2010). The cognitive interview: A meta-analytic review and study space analysis of the past 25 years. *Psychology, Public Policy, and Law, 16*, 340–372. https://doi.org/10.1037/a0020518.

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*, 184–211. http://doi.org/10.1207/s15327957pspr0903_1.

Paggio, P., & Navarretta, C. (2013). Head movements, facial expressions and feedback in conversations: Empirical evidence from Danish multimodal data. *Journal on Multimodal User Interfaces, 7*, 29–37.

Parkinson, B., Briner, R. B., Reynolds, S., & Totterdell, P. (1995). Time frames for mood: Relations between momentary and generalized ratings of affect. *Personality and Social Psychology Bulletin, 21*, 331–339. http://doi.org/10.1177/0146167295214003.

Purvis, J. A., Dabbs, J. M., Jr., & Hopper, C. H. (1984). The" opener" skilled user of facial expression and speech pattern. *Personality and Social Psychology Bulletin, 10*, 61–66.

Raposo, A., Moss, H. E., Stamatakis, E. A., & Tyler, L. K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia, 47*, 388–396. http://doi.org/10.1016/j.neuropsychologia.2008.09.017.

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications.* London: John Wiley & Sons.

Spitzberg, B. H. (2006). Preliminary development of a model and measure of computer-mediated communication competence. *Journal of Computer-Mediated Communication, 11*, 629–666. http://doi.org/10.1111/j.1083-6101.2006.00030.x.

Straker, L. M., Coleman, J., Skoss, R., Maslen, B. A., Burgess-Limerick, R., & Pollock, C. M. (2008). A comparison of posture and muscle activity during tablet computer, desktop computer and paper use by young children. *Ergonomics, 51*, 540–555. http://doi.org/10.1080/00140130701711000.

The Skype Team. (2017). *One billion downloads of Skype for Android – thank you!*. Available at: https://blogs.skype.com/news/2017/10/25/one-billion-downloads-of-skype-for-android-thank-you/, Accessed date: 31 January 2019.

Tidwell, L. C., & Walther, J. B. (2002). Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human Communication Research, 28*, 317–348. http://doi.org/10.1111/j.1468-2958.2002.tb00811.x.

Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23*, 3–43. http://doi.org/10.1177/009365096023001001.

Walther, J. B., Van Der Heide, B., Hamel, L. M., & Shulman, H. C. (2009). Self-generated versus other-generated statements and impressions in computer-mediated communication: A test of warranting theory using Facebook. *Communication Research, 36*, 229–253. https://doi.org/10.1177/0093650208330251.