

The impact of head movements on user involvement in mediated interaction

Abstract

Communication takes place not only through speech, but also through gestures such as facial expressions, gaze, head movements, hand movements and body posture. Although developing rapidly, current communication platforms do not facilitate the types of behaviour we believe are needed to fully support non-verbal communication and make interactions more engaging and efficient. In this paper, we decided to focus our research specifically on the head rather than any other body part as it is a rich source of information for speech-related movement. Thus we aim in this study to investigate the value of incorporating head movements into the use of telepresence robots as communication platforms; by means of investigating a system that manually reproduces head movement as closely as possible. The essential quantitative results revealed no significant differences on any of the measures we used. However, the qualitative information from the experiment indicates of further research will be useful in this area. These findings suggest that an enclosed body language are required for a real-time communication beside the head nodding.

Keywords: *engagement, nonverbal behaviours, head movements, face-to-face interaction, telepresence robot.*

1 Introduction

Communication networks are social environments. Thus, analysing their performance must cover both engineering criteria, such as telecommunication bandwidth, and social criteria such as group communication, what we might call social presence. Basically social presence is the sense of facing remote people in a shared space, or in other words the degree of resembling face-to-face interaction (De Greef & Ijsselstein, 2001). Successive generations of emerging networked interfaces have been designed to mediate remote social communication.

One of the early theories that drove the recent research about social presence can be traced back to the end of the 1960s; Mehrabian's (1968) concept of immediacy, which he defined as "those communication behaviours that enhance closeness to and nonverbal interaction with another". The emphasis on interactive behaviour is followed by a more recent definition, as in Palmer's (1995) definition of social presence which includes "effectively negotiate(ing) a relationship through an interdependent, multichannel exchange of behaviours" and Heeter's (1992) definition which emphasises reaction and interactivity. These definitions propose that an increase in the social presence is related to an increase in the engagement that includes non-verbal cues and signals which take place normally within face-to-face communication, such as gaze, head and hand gestures etc. Therefore, increasing social presence is crucial in designing an effective system and has become

an important goal for many research projects in the area of telepresence robots (TP).

2 Background

Face-to-face communication takes place not only through speech, but also by means of several nonverbal behaviours such as facial expressions, gaze, head movements, hand movements and body posture. These behaviours do not play a direct part in the articulation of speech, but they are produced to provide complementary information to speech, sustaining and regulating interaction and finally providing continuity, serving important syntactic and prosodic functions such as the stress and intonation patterns of an utterance (Burgoon & Saine, 1978).

Until recently, video conferences and instant messaging have been the only ways to convey nonverbal information between people at a distance. However, the problem with these types of systems is that they cannot provide the full breadth of visible behaviour content that we exchange in our daily collocated interaction. This is because they only display the image of the person's head and shoulders, which cannot fully convey their presence (Sirkin, Ju, & Cutkosky, 2012). Though the video conferencing system overcomes some of these issues by capturing a greater range of behaviour, it is still the case that the presence of the person is liable to be inhibited because of the lack of capacity for movement. The person's image is limited by the screen and they become literally and figuratively two dimensional. They cannot move, touch, refer to or even manipulate objects around the physical space where the team is located. The telepresence robotic system is meant to solve the issues with the chat and video conferencing systems; as it offers a means to connect to a remote location via traditional telepresence with the added value of moving

and actuating in that location. However, a major concern with such technologies is that they are often lacking in the social presence of the human, although the human has a physical presence through the robot.

Recent works discuss guidelines for increasing social acceptability of a robot appearance and interface aspects (Mollahosseini et al., 2014; Li, Rau, & Li, 2010; Schillaci, Bodiroža, & Hafner, 2013; Arras & Cerqui, 2005 and Schröder et al., 2012). Whereas (Lee & Takayama, 2011 and Tsui, Desai, Yanco, & Uhlik, 2011) explored the perspective of co-workers by investigating the bystander's impression of a TP robot system placed in an office, and discussed guidelines for increasing social acceptability. In addition, there has been a significant amount of these studies were related to the sense of presence generally and social presence of users.

However, the focus of most of these studies was on the function or utility of the robot, ignoring the relationships between the user characteristics and the feelings towards the robot. Based on this concept and in order to have more effective and immediate interactions, a robot needs to simultaneously exhibit competent behaviour, convey attention and intentionality, and handle social interaction. Most of current telepresence robots which have been developed for commercial purposes are mainly focusing on the physical capabilities of the robots, ignoring the vital importance of the full breadth of engagement behaviour content humans experience in our daily life (Bamoallem, Wodehouse, & Mair, 2014). Thus, exhibition of naturalistic behaviour and appropriate emotions by the robot is the main core for an effective system (Doherty-Sneddon et al., 1997 and Bates, 1994). Based on this our system has the ability to convey their operators' head movements in order to

make remote interactions more present, more engaging.

2.1 Conceptual framework

In this study we proposed to focus on measuring engagement within conversational behaviours as they provide significant evidence of connection between the participants as supported by different studies Sidner, Lee, Kidd, Lesh, & Rich, 2005 and Cassell, 2000.

2.1.1 Conversational involvement

Several studies have looked at how the structure of conversation changes with communication mode (Cook & Lalljee, 1972 and Rosenfeld, 1978). Measures of conversation structure include some nonverbal behaviours such as the number and length of speaker exchanges and the number of pauses and interruptions in free speech (U. Hadar, T. Steiner, & F. C. Rose, 1985). A common finding in the literature is that face-to-face conversations result in more turns, shorter lengths of turn and more interruptions than audio-only or video-mediated dialogues (O'Conaill, Whittaker, & Wilbur, 1993; Sellen, 1995). The interpretation of these findings has been that nonverbal behaviour or visual signals are important, thus face-to-face communication is less formal, with more interruptions (simultaneous speech) and fewer formal handovers of turns (Beattie & Barnard, 1979; Ellis & Beattie, 1986 and Rutter & Stephenson, 1977). The underlying assumption behind these differences may result from technical limitations or a lack of nonverbal behaviour as (Short, Williams, & Christie, 1976) highlighted that the lack of head turning and directional gaze in many video mediated communication (VMC) systems may affect turn-taking behaviour.

Hence, it is to be expected that systems that are careful to preserve important non-verbal

cues will make a difference in overall interaction behaviour. A related question is the extent to which the media difference will influence participants' interaction involvement. Recent studies suggest that socially expressive media can significantly improve viewers' interpretation of the action, and it can be seen to be more engaging and likable than a static one (Adalgeirsson & Breazeal, 2010; Sirkin & Ju, 2012; Mutlu, Yamaoka, Kanda, Ishiguro, & Hagita, 2009). It is conceivable that the media could have some observable influence on cognitive activity such as interaction involvement.

Conversational involvement has not attracted much scholarly attention and few studies up to now have made a deep level of investigation into this field to extend the way in which involvement is measured. This is particularly the case in conversational involvement in conjunction with a telepresence robot.

2.1.2 Measurement of conversation involvement

Observational rating

The interest in interpersonal interaction developed different observational ratings systems of nonverbal involvement, the most known and established ones are (Coker & BURGOON, 1987; (Laura Knarr Guerrero, 1994 and Laura K Guerrero, 2005). The Guerrero system was developed using Coker and Burgoon items as a guide to measure specific involvement behaviours that could be rated by coders. For this we have adopted the use of this system in our study as it the most up-to-date version of these rating systems.

The Guerrero observational rating system (Laura K Guerrero, 2005) was developed to rate behavioural cues of involvement in human dyadic interaction, and thus, we argue it can also be used with human mediated

interaction. The system is comprised of six scales that are necessary for measurement of involvement. These are; immediacy, expressiveness, altercentrism, interaction management, composure and positive affect. If an individual is showing a greater immediacy, greater expressiveness, better interaction management, more altercentrism, lack of concern about others, greater negative arousal and more positive behaviours, it can be reasonably assumed that that individual is highly integrated in their feelings, thoughts and experiences with the on-going interaction - a highly involved individual. (1) *Immediacy* dimension behaviours measure the physical proximity between two individuals; (2) *Expressiveness* dimension behaviours communicate the level of energy, activity, and enthusiasm toward conversation partner; (3) *Altercentrism* dimension behaviours reflect the degree of focus on the conversation partner during interaction; (4) *Interaction management* dimension behaviours support a smooth flow of conversation; (5) *Composure dimension* behaviours reflect an absence of nervous body movement or the presence of confidence; and (6) *Positive affect* dimension behaviours include smiling, laughing and other behaviours that reflect good feelings about the interaction and the partner.

We believe that the Guerrero (Laura K Guerrero, 2005) system is uniquely suited for measuring a wide range of nonverbal and verbal indicators to determine the degree to which an individual is actively involved in a real-time conversation, as in our case. Another study (Norris, Weger, Bullinger, & Bowers, 2014) added that although the system was designed to include six different dimensions, it could be altered to use particular dimensions according to the focus of the research without impacting measurement reliability and validity. This can be seen as an advantage in our present study

where the immediacy dimension is not applicable because of the nature of the game technology used in our research; there is no physical contact between participants. Hence we examined gaze within immediacy, nods within altercentrism, smiling within positive effects and lack of random movement within composure.

Based on the previous section the following hypotheses were generated

Simulating head nodding in a telepresence conversation will increase involvement communicated by more eye gaze, more altercentrism seen by more nodding, positive feeling between both sites can be seen by more smiling and laughing, less number of answers and questions, and finally lack of random movement by composure.

Self-reporting

Cegala (1982) conduct one of the early studies about interaction involvement identified three dimensions of involvement: responsiveness (that is, mental alertness to the situation), perceptiveness (that is, ability to make attributions about one's or others' behaviour), and attentiveness (that is, awareness of factors impacting interaction).

This questionnaire was chosen because it is widely regarded as the most relevant and comprehensive instrument available to assess individual's personal tendencies relative to involvement in communication settings And also used extensively in literature (Duran & Kelly, 1988; Sidelinger, Ayash, Godorhazy, & Tibbles, 2008 and Norris, et al., 2014)

As mentioned earlier, IIS consists of attentiveness, responsiveness and perceptiveness. Although each of these factors can be examined independently, an overall score can be gleaned in an effort to assess one's overall tendencies toward interaction. To our knowledge, few studies

have made a comparison between conversation involvement measures (self-reporting and observational ratings measures), especially when a cross-situation comparison is made between two contrasting settings such as the ones under study here. Therefore, it is of interest to compare conversational involvement of participants (objective measure) and their subjective ratings of satisfaction about involvement in the task. Based on this, we hypothesis that a high score in conversational involvement (Objective) will be associated with a high score on an interaction involvement scale (Subjective).

3 Study overview

As we have previously stated, telepresence in the context of this application means replacement of human presence with a robot, which is operated by a human driver from a location at a distance. Increasing presence is crucial in designing an effective robot system, thus it becomes the goal of most research projects in the area of telepresence robots. Our research aimed to investigate the influence that telepresence has on communication, by examining sense of engagement as a part of social presence which can be improved further by adding more non-verbal cues.

The study of nonverbal communication during conversations, particularly the study of head movement, is an extremely rewarding field. They effectively give real-time listener reactions, and form part of the feedback loop that we all rely on to tell us how effective our communication is being (U. Hadar, T. J. Steiner, & F. C. Rose, 1985). Thus, this study aim to examine the face channel which contains some of the nonverbal cues such as movements and expressions, focusing on

head movement which includes nodding and head orientation.

4 Methods

4.1.1 Study overview

Our overall position is that making telepresence systems socially expressive by affording them the ability to convey their operators' non-verbal behaviours such as gesture and posture can make remote interactions more present, more engaging.

To test this claim, we design an experiment where we could evaluate the experience of the subject when interacting with a socially expressive system. As we specifically wanted to learn about how head movements affect the collaborator's experience, we decided to run a within subject study with two conditions to give a better understanding of how head movement might affect the conversational engagement (Figure, 1):

1. The simulated condition: Participants interacted through a video call where face and head movements represented on the screen and the screen replicated the head movements of the person on the screen.
2. The video mediated condition: The screen is in neutral and still (non-moving) poses during the whole interaction and participants interacted through a video call where face and head movements represented on the screen only.

Our experimental design involves one manipulated independent variable, which is whether or not the screen produces the head movement and physically nods. This variable was generated by two different scenarios and conditions as previously stated; video

mediated on-screen movements, and video mediated on-screen and in-space movements. The dependent variables involve both self-report and observation measurements during the completion of a task.

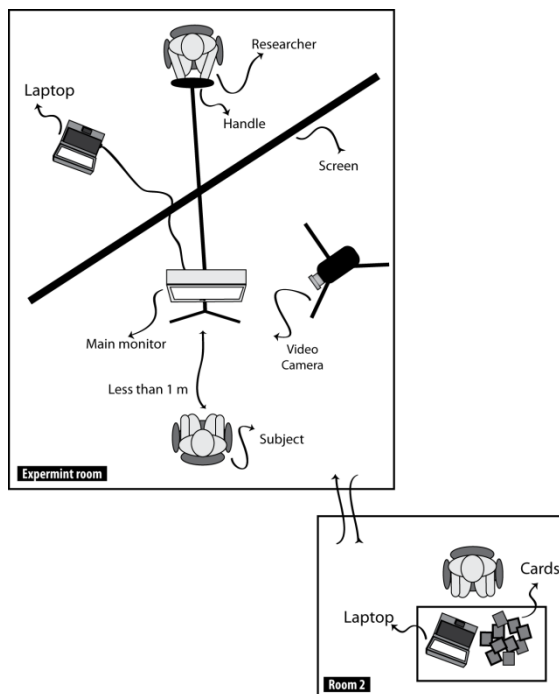


Figure 1 Experiment Design

4.1.2 Experimental task

Although it has been supported by different studies the important effects that visual signals have on communication outcomes, a conflict in the task outcomes has been highlighted. Some studies found that for problem-solving tasks, face-to-face and audio-only interactions do not differ in terms of task outcome (Chapanis, Ochsman, Parrish, & Weeks, 1972 and Williams, 1977). In contrast, with design tasks and social tasks involving negotiation or conflict resolution, performance was better in face-to-face or mediated communication than audio-only (Olson, Olson, & Meader, 1995 and Short, et al., 1976). Therefore we decided to start with a study done by Mutlu (2009) who examined communication and task performance in human-robot communication (VMC) using guessing game. We chose to use the guessing

game task because it has the following characteristics which are beneficial to our evaluation as:

- It is intended to allow study of the dialogue generated during the task
- It sparks an active conversation between both sides
- It takes under about thirty minutes to finish

However, the format was adapted according to our context of use:

- In our experiment we gave the picker the opportunity to explain the item he/she picked as we are looking to have a dynamic interaction between both sides which it will be difficult to achieve if we use the original game design.
- In the original game design both participant will have all the items in front of them as they are facing each other. The situation was not part of our experiment design and therefore we place the items in front of the picker only.

Experiment design

We devised an experimental task in which two participants were to play the guessing game. In the game, one of the players (the Picker) would choose an item - without identifying it to the other player - from among eighteen items (printed onto A5 size cards) placed in front of him/her. The other player (the Guesser) would try to guess which item the Picker had chosen by asking the Picker a set of questions that could be answered with "Yes" or "No". However, the Picker could give the Guesser some hints if they wanted to do so. By this method, both sides would have an equal chance to interact and exchange words with each other (Figure 2)

For the purposes of the task, we were careful to select items that would be in common use in the UK, in a balanced set of materials, shapes sizes and colours. Both participants were provided with detailed instructions on how the game should be played.

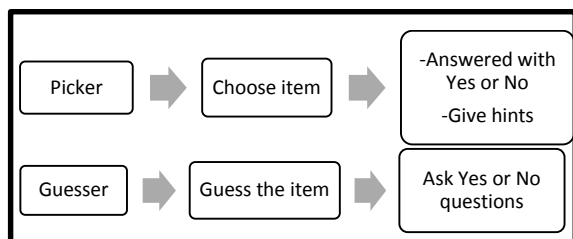


Figure 2 Experiment Procedure

4.1.3 Participants

The participants took part in the experiments in pairs; with each experiment involving up to 13 pairs, making a total of 26 subjects. 3 pairs were from different gender, and the rest were from the same gender (3 female pairs and 7 male pairs). The total number of experiments was 26, 13 with each conditions. Each pair consisted of a Picker and a Guesser. The ages of the subjects varied between 18 and 29; and their subjects represented a variety of university majors.

4.1.4 Technical equipment

Based on the previous findings, we developed a model which is a monitor on a stand. The researcher will be able to replicate the head movement for the person on the screen manually by using Puppeteer mechanism in the monitor's stand which is basically a way to inanimate objects to life by mimicking actions with simple hand movement. The researcher replicated the basic movements of the head. Those are nod -up and down movements- which used to show our agreement with what is saying and shake -lift and right movement- to express disapproval and negation.

Hardware

- 12" screen: used in the study for all participants in the last condition; mounted behind a black curtain.
- Video recordings: a digital video camera recorded the Guesser side in all conditions, and was placed 1 metre away from the speaker in the recording studio. This was used to record the dialogue and the interactions of the participants and covered three subjects (screen, body and face and the task document).
- Laptop: 1 laptop used for video call conditions (first condition) from the picker side.
- High-definition video webcams were mounted at the top of the screens for both conditions.

Software

Skype application was used for the video call.

4.1.5 Procedure

The purpose of the experiment was to gather real-time data under two different conditions: video mediated and video mediated with in-space screen movement.

At the beginning of the session, we provided the participants with a brief description of the purpose and procedure of the experiment, but we deliberately concealed the primary purpose of the experiment. We then asked the participants to fill in a pre-experiment questionnaire on their affective state and some background data, such as gender and age etc.

The participants were given an introduction on how the experiments would operate; and they were instructed how to play the Guessing game, allocated a role as a Picker or Guesser, and told what they needed to do. After that, we took the participants into the experiment room to start the task.

The sessions took place in two rooms, one equipped with laptop only, where the Pickers sat. The second room was the experimental room for the Guesser where we place a modified rig for the 12" monitor and video camera. We focused only on the Guesser side as it communicates more than the Picker side. The Guesser has to ask questions, give explanation and give answers, where the Picker side has only to say "Yes" or "No" and give hints. The procedures were the same under both different conditions; the researcher assisted each subject with placement of the webcam image and to ensure that the subjects were comfortable and then they were asked to begin the test. Each pair completed two Guessing games under different conditions which lasted approximately 15 min. Guessers sat in front of the modified rig that held the 12" monitor with a small Web camera affixed to the top. Following the game, participants were given a 5 to 10 minute break to complete the Interaction Involvement Scale (IIS) questionnaire (Section 4.2.1). This questionnaire took about 5 min to complete.

First Condition

For the mediated conditions, subjects were not able to see each other physically, but could work collaboratively together and were able to hear and talk to each other without wearing headsets. A screen was erected between them, adjusted to block any direct view of one another's faces, but they could see one another's faces through the webcams placed on top of each screen. Both sides had a camera set up to record their interactions.

Second Condition

Similar to the first condition; however, the experimenter took a seat beside the screen displaying the head movements of one side,

and reproduced those movements on the other side's screen as closely as possible.

4.2 Measurement

Once we had identified the factors, we needed to decide how we could assess them. As previously stated, Biocca & Harms (2002) explained that we needed to go beyond technology assessments, and more into the realms of psychology and sociology. We therefore decided that the project would involve a variety of means of information gathering and experiments using questionnaires and video recordings of practical interaction experiments.

One of the challenges in the work was to decide how to measure the impact of the in-space movement on the participant's interaction. Our experimental design involved one manipulated independent variable; whether the screen produced movement or not. The dependent variables involved *objective* and *subjective* measurements.

4.2.1 Objective

We videotaped the sessions for both conditions and reviewed the recordings, transcribing them under a number of headings as discussed earlier in this report (section 2.1.2), to determine interaction behaviour during the sessions.

Behavioural observations

We measured participants' task performance through capturing the number of questions they asked and the answers they gave, to identify the Picker's choice. We also measured the involvement score for all Guessers for the two conditions. As discussed earlier in this report (section 2.1.2). All sessions were videotaped to support the analysis of the objective measures.

4.2.2 Subjective

Subjective measures were used to obtain ratings of the subjects' satisfaction in their involvement in the game using the Interaction Involvement scale (IIS) and to evaluate the appropriateness of the movements. This was used to answer H7.

Involvement

Upon completion of the task the subjects were required to complete a post-test questionnaire which is version of The Interaction Involvement Scale (IIS) as discussed previously in (section 2.1.2).

Response options for each IIS item range from 1 (not at all like me) to 5 (very much like me). Responses to some items were reverse coded and high scores refer to high interaction involvement with communication.

Effectiveness of movements

Part of this questionnaire was a version derived from (Sidner, Lee, & Kidd, 2005) (4 items). This measure was given to the participants at the end of the second session, as it measured the appropriateness of the head movements. Response options for each item ranged from 1 (Disagree) to 5 (Agree).

In addition we used open questions to gather information about our subjects' experience.

4.2.3 Data analysis

Mixed methods were the choosing for this exploratory experiment to better understand a specific issue of the quantitative study using the quantitative ones base on the participants views (Creswell & Plano Clark; 2007; O Cathain, 2010).

4.2.4 Preparation of involvement rating data for analysis

Behaviours were examined prior to testing the study hypotheses to search for any sign of systematic patterns consistent with errors or

missing data. Based on this examination we found three different systematic patterns in our data. The first case had little or no incidences of head nods; adopters and twisting behaviours observed. We also faced a problem of missing data from the game scenario, some recording did not provide a full view of the player either because of the player covering the head with hands or scarf, or because the player leaned too far forward situating themselves partially or completely out of view of the video camera, resulting in no rateable nonverbal behaviour. As a result, we prevented these items from being used in further analysis altercentrism and composure.

The continuous kind of behaviours by the subjects such as smiling and eye gaze were counted simply using electronic stop watches, whereas the rest of the behaviours were counted by the numbers of times they exhibited them, such as the number of questions and answers, nods and adopters or twisting behaviours. We also noted that interaction length varied between 2 to 7 min due to player style and personal self-confidence; we found that shy, taciturn player's interactions were shorter than talkative players. This could have caused an issue with our data, as the score for the number of behaviours would have varied as well. To overcome this issue, we converted the score to the number of behaviours per 1 minute by dividing the score by the time taken.

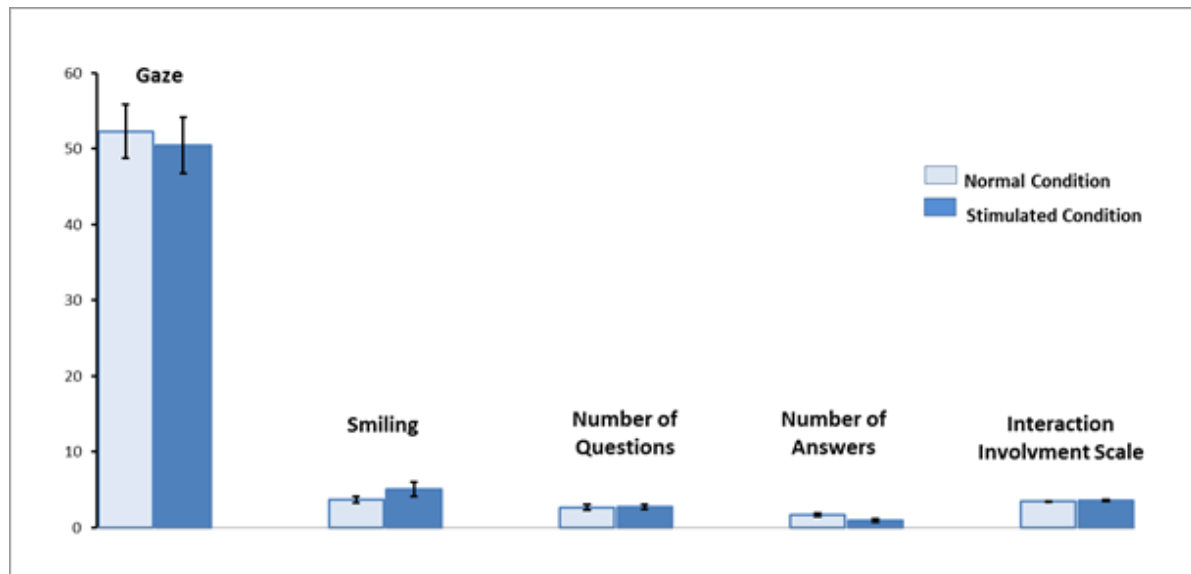


Figure 3 t-test Finding between Normal and Stimulated Conditions

5 Results

In general, this investigation aimed to shed light on the influence that head movement has on communication. Using Paired t-test we analysed the gaze, smiling and number of questions and answers between the two conditions. We hypothesised that subjects in the simulated condition scored highly in conversational involvement, communicated by better gaze, smiling and number of questions and answers. The results revealed no significant differences on any of the measures we used between the two conditions ($p > .05$), as we can see from figure (3).

Smiling and number of questions in the simulated setting reported overall higher scores ($M = 3.3$; $SD = 4.72$) for smiling and ($M = 3.21$; $SD = 1.58$) for number of questions than in normal setting ($M = 2.43$; $SD = 2.07$) and ($M = 3.15$; $SD = 1.71$). Whereas, eye gaze and number of answers in the normal setting reported an overall higher score ($M = 46.2$; $SD = 18.32$) for gaze and for number of answers ($M = 1.47$; $SD = 1.26$) than in the simulated setting ($M = 46.17$; $SD = 18.96$) for gaze and ($M = 1.19$; $SD = 1.01$) for number of

answers. Also, a t test was run on IIS scores H7, the results showed no significant difference between the group means ($p = .582$).

Despite these result, our qualitative data showed that about 90% of our subjects were satisfied with the movement in the second condition as we mentioned earlier in the result section.

In addition to this, compared subject satisfaction with the movement using the, which resulted in fluctuation between them, minimum one was less than 2 and maximum was almost 5.

6 Discussion

A review of the literature on interaction involvement and communication behaviour indicates that nonverbal behaviours play an important role in message production and in involvement in a variety of different situations which inspired this study of possible similar findings. However, we did not find any improvements in the respondents with the use of video call and video simulated call. These results do not support the results from literature review; this might be for different

reasons. With the Interaction Involvement Scale the results did not show any statistically significant improvement. Although the present study's sample was representative of college-aged students, and covered an international population, we observed that some of the students were asking for explanations when they answered the questionnaire. We can argue that it is possible that respondents did not accurately answer our questions as many of them were international students and did not fully understand the questions. Secondly, the amounts of time the participants took to complete the task ranged from 2 to 7 min as they have to complete maximum of 4 guessing. This in our opinion was not a sufficiently large difference to have a noticeable impact on their communication patterns. An alternative explanation is that we might argue that interaction involvement is best defined as a personal trait, which is less likely to be influenced by the context in which the communication is taking place, and more to do with the personal traits of the participants. Villaume, Cegala (1988) stated that people who are highly involved understand the relational aspect of conversations rather than the content of the conversation, so a highly engaged person might ignore the channel and place emphasis on the relational signals. Other rationales may exist in the motivation for interpersonal communications and also factors such as interaction involvement.

In respect to the specific behaviour ratings related to involvement, there are different explanations for our findings. Some subjects reported that the visual and the audio signals were not particularly good, as there was sometimes a delay; this might on the surface appear plausible and explain some of the results we obtained. It is possible that the technology we used resulted in a less than

ideal image transmission, due to limited bandwidth and technology constraints. As a result, resolution or frame rate may have suffered, or signals may have been delayed. A good way to test this supposition would be the use of technology which made it possible for video-mediated conversation to look as similar to a face-to-face conversation as possible (high quality image resolution, no transmission delays), simulating eye-to-eye contact. In fact, Doherty-Sneddon, (Doherty-Sneddon, et al., 1997) found that video-mediated conversations tend to be more formal, with fewer interruptions and longer utterances, especially when there is asynchrony or delay in visual and audio signals.

Another work by O'Malley et,al (O'Malley, Langton, Anderson, Doherty-Sneddon, & Bruce, 1996) explained that some studies cited have involved fairly open-ended discussions and debates. Simultaneous speech and interruptions tend to occur in less formal and more spontaneous circumstances. In our task, the shorter turns and high level of interruptions may have been indicative of the artificiality of the situation causing awkwardness among the participants, and difficulty in smooth turn-taking. We presume from this, that the more conversation cues there are, the more participants will involve in the conversations. In other words, people tend to interrupt each other, when there is a problem in regulating a conversations for example overlapping speech between participants. We also, believe the similarities between the two conditions in the quantitative data in the present study may also be related to the effect of the size of the video image. Because of the technology we were using, we couldn't just shrink the video image to only give a view of the head and neck. Instead, a view of the head and shoulders, down to the elbow was present in

both conditions. As result, subjects were mainly using visual cues from the face (e.g. gaze, expression, lip-movements) but also to a degree global cues such as posture and gesture (e.g. shrugging of shoulders). Therefore some users reported that the simulated condition caused distraction during the interaction. An explanation for this is due to the size of the video image, as the movement was meant to be for the face image only, not the shoulders as well.

On the other hand, a good part of the participants reported that the simulated condition helped them in understanding and communicating much better with their partners and the movement helped in making the conversation more efficient which as a positive findings support our hypothesis. One subject said:

“In the second condition I think the screen movement helped me to interact in a more effective and efficient manner, I was reminded to keep my eyes on the screen and my conversation partner when the screen moves”.

“I prefer the conversation in the second condition more, as it helped me to better understand what my partner intends to express”

“I like when the screen replicating the movements, in which provides emphasis more about the partner’s expression”

7 Concluding remarks and future work

The aim of this study was to explore the potential of replicating the head movements that support engagement and feedback functions between people, which could be useful in the development of a model of human gestures to implement in a

telepresence platform. The results did not show any evidence of the possibility of identifying any improvement by incorporating head movements.

Despite the limitations, our findings at least highlight important implication for future research. These findings suggest that face-to-face interaction is complex in its own right as it includes various behaviours that help in maintaining the connection between two people. Thus it will be difficult to find any significant result if we only focus on one of these behaviours. In general, it should be noted that real-time communication requires more than verbal communication, facial expressions and head nodding. It is important to complement them with other types of nonverbal behaviour such as posture. Also, it is important to highlight that head movements as displayed in conversation are more dynamic than we could achieve with our system. The head moves constantly while talking, and if the head is still, it tends to be interpreted as a pause or as listening (Pittenger, Hockett, & Danehy, 1960; Birdwhistell, 1983). Also, Speech-related head movements range in amplitude between 170 and zero, with a range of speeds (Hadar, Steiner, Grant, & Rose, 1983). Thus, we needed to produce real-time dynamic communication in order to get a significant result which was not the case with our experiment. Perhaps participants responded less quickly to the movement because the movements did not convey the same sense of urgency. More accurate reproductions and simulations of head movements might give objective improvement in remote communication.

Although our measurement and analysis methods used in this study appear to be a practical way to acquire data on the magnitude and length of a variety of gestures, we also believe that available technology can

be used to further measure other specifics and synchronizing the head movements such as eye gaze, using an eye gaze tracking system, webcam as video recorder and most importantly head tracking system to transfer the head movement in real-time to the movement of the screen. Thus, we plan to produce precise on-screen movement by synchronize the on-screen movement with the head movement of the participants which was one of the main issues we faced in this study. Providing full details about how we plan to resolve this synchronization problem will be our next step.

References

- Adalgeirsson, S. O., & Breazeal, C. (2010). *MeBot: a robotic platform for socially embodied presence*. Paper presented at the Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction.
- Arras, K. O., & Cerqui, D. (2005). *Do we want to share our lives and bodies with robots? A 2000 people survey*.
- Bamoallem, B., Wodehouse, A., & Mair, G. (2014). *DESIGN FOR AN OPTIMAL SOCIAL PRESENCE EXPERIENCE WHEN USING TELEPRESENCE ROBOTS*. Paper presented at the DS 77: Proceedings of the DESIGN 2014 13th International Design Conference.
- Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37(7), 122-125.
- Beattie, G. W., & Barnard, P. (1979). The temporal structure of natural telephone conversations (directory enquiry calls). *Linguistics*, 17(3-4), 213-230.
- Biocca, F., & Harms, C. (2002). Defining and measuring social presence: Contribution to the networked minds theory and measure. *Proceedings of PRESENCE*, 2002, 1-36.
- Birdwhistell, R. L. (1983). Background to kinesics. *Etc.*, 40(3), 352-361.
- Burgoon, J. K., & Saine, T. P. (1978). *The Unspoken Dialogue: An Introduction to Nonverbal Communication*: Houghton Mifflin School.
- Cassell, J. (2000). Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. *Embodied conversational agents*, 1-27.
- Cegala, D. J., Savage, G. T., Brunner, C. C., & Conrad, A. B. (1982). An elaboration of the meaning of interaction involvement: Toward the development of a theoretical concept. *Communications Monographs*, 49(4), 229-248.
- Chapanis, A., Ochsman, R. B., Parrish, R. N., & Weeks, G. D. (1972). Studies in interactive communication: I. The effects of four communication modes on the behavior of teams during cooperative problem-solving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 14(6), 487-509.
- Coker, D. A., & BURGOON, J. (1987). The nature of conversational involvement and nonverbal encoding patterns. *Human Communication Research*, 13(4), 463-494.
- Cook, M., & Lalljee, M. G. (1972). Verbal substitutes for visual signals in interaction. *Semiotica*, 6(3), 212-221.
- De Greef, P., & Ijsselsteijn, W. A. (2001). Social presence in a home tele-application. *CyberPsychology & Behavior*, 4(2), 307-315.
- Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., & Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied*, 3(2), 105.
- Duran, R. L., & Kelly, L. (1988). An investigation into the cognitive domain of competence II: The relationship between communicative

- competence and interaction involvement. *Communication Research Reports*, 5(1), 91-96.
- Ellis, A. W., & Beattie, G. (1986). *The psychology of language and communication*: Weidenfeld and Nicolson.
- Guerrero, L. K. (1994). An application of attachment theory to relational messages and nonverbal involvement behaviors in romantic relationships.
- Guerrero, L. K. (2005). Observer ratings of nonverbal involvement and immediacy. *The sourcebook of nonverbal measures: Going beyond words*, 221-235.
- Hadar, U., Steiner, T., Grant, E., & Rose, F. C. (1983). Kinematics of head movements accompanying speech during conversation. *Human Movement Science*, 2(1), 35-46.
- Hadar, U., Steiner, T., & Rose, F. C. (1985). Head movement during listening turns in conversation. *Journal of Nonverbal Behavior*, 9(4), 214-228.
- Hadar, U., Steiner, T. J., & Rose, F. C. (1985). Head movement during listening turns in conversation. *Journal of Nonverbal Behavior*, 9(4), 214-228.
- Heeter, C. (1992). Being there: The Subjective Experience of Presence. *Presence: Teleoper. Virtual Environment*, 1(MIT Press).
- Lee, M. K., & Takayama, L. (2011). *Now, I have a body: Uses and social norms for mobile remote presence in the workplace*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Li, D., Rau, P. L. P., & Li, Y. (2010). A cross-cultural study: Effect of robot appearance and task. *International Journal of Social Robotics*, 2(2), 175-186.
- Mehrabian, A. (1968). Some referents and measures of nonverbal behavior. *Behavior Research Methods & Instrumentation*, 1(6), 203-207.
- Mollahosseini, A., Graitzer, G., Borts, E., Conyers, S., Voyles, R. M., Cole, R., et al. (2014). *ExpressionBot: An emotive lifelike robotic face for face-to-face communication*. Paper presented at the Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on.
- Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2009). *Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior*. Paper presented at the Proceedings of the 4th ACM/IEEE international conference on Human robot interaction.
- Norris, A. E., Weger, H., Bullinger, C., & Bowers, A. (2014). Quantifying engagement: Measuring player involvement in human–avatar interactions. *Computers in human behavior*, 34, 1-11.
- O'Conaill, B., Whittaker, S., & Wilbur, S. (1993). Conversations over video conferences: An evaluation of the spoken aspects of video-mediated communication. *Human-computer interaction*, 8(4), 389-428.
- O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G., & Bruce, V. (1996). Comparison of face-to-face and video-mediated interaction. *Interacting with computers*, 8(2), 177-192.
- Olson, J. S., Olson, G. M., & Meader, D. K. (1995). *What mix of video and audio is useful for small groups doing remote real-time design work?* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Palmer, M. T. (1995). Interpersonal communication and virtual reality: Mediating interpersonal relationships. *Communication in the age of virtual reality*, 277-299.
- Pittenger, R. E., Hockett, C. F., & Danehy, J. J. (1960). The first five minutes: A sample of microscopic interview analysis.
- Rosenfeld, H. M. (1978). Conversational control functions of nonverbal

- behavior. *Nonverbal behavior and communication*, 291-328.
- Rutter, D. R., & Stephenson, G. M. (1977). The role of visual communication in synchronising conversation. *European Journal of Social Psychology*, 7(1), 29-37.
- Schillaci, G., Bodiroža, S., & Hafner, V. V. (2013). Evaluating the effect of saliency detection and attention manipulation in human-robot interaction. *International Journal of Social Robotics*, 5(1), 139-152.
- Schröder, M., Bevacqua, E., Cowie, R., Eyben, F., Gunes, H., Heylen, D., et al. (2012). Building autonomous sensitive artificial listeners. *Affective Computing, IEEE Transactions on*, 3(2), 165-183.
- Sellen, A. J. (1995). Remote conversations: The effects of mediating talk with technology. *Human-computer interaction*, 10(4), 401-444.
- Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications.
- Sidelinger, R. J., Ayash, G., Godorhazy, A., & Tibbles, D. (2008). Couples go online: Relational maintenance behaviors and relational characteristics use in dating relationships. *Human Communication A Journal of the Pacific and Asian Communication Association*, 11(3), 333.
- Sidner, C. L., Lee, C., & Kidd, C. (2005). *Engagement during dialogues with robots*. Paper presented at the AAAI spring symposia.
- Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., & Rich, C. (2005). Explorations in engagement for humans and robots. *Artificial Intelligence*, 166(1), 140-164.
- Sirkin, D., & Ju, W. (2012). *Consistency in physical and on-screen action improves perceptions of telepresence robots*. Paper presented at the Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction.
- Sirkin, D., Ju, W., & Cutkosky, M. (2012). Communicating Meaning and Role in Distributed Design Collaboration: How Crowdsourced Users Help Inform the Design of Telepresence Robotics *Design Thinking Research* (pp. 173-187): Springer.
- Tsui, K. M., Desai, M., Yanco, H., & Uhlik, C. (2011). *Exploring use cases for telepresence robots*. Paper presented at the Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference on.
- Villaume, W. A., & Cegala, D. J. (1988). Interaction involvement and discourse strategies: The patterned use of cohesive devices in conversation. *Communications Monographs*, 55(1), 22-40.
- Williams, E. (1977). Experimental comparisons of face-to-face and mediated communication: A review. *Psychological Bulletin*, 84(5), 963.