

BRING YOU CLOSER TO ME

*Does visually co-located video-mediated-communication
draw more attention to the conversation?*

Pinhsien Lee

Supervisors: Maarten H. Lamers, Marcello Gómez Maureira

*leepinhsien@gmail.com | Media Technology MSc Graduation Thesis
Leiden University*

Bring you closer to me

Does visually co-located video-mediated-communication draw more attention to the conversation?

Pinhsien Lee

Graduation Thesis, July 2018

Media Technology MSc program, Leiden University

Supervisors: Maarten H. Lamers, Marcello Gómez Maureira

leepinhsien@gmail.com

Abstract:

This research focused on the visual factor of the background in video-mediated communication (VMC). Regarding its influence on people's attention level and the feelings of closeness. Subjects in the experimental group had an online simulated VMC with the speaker whose background was visually co-located with the subjects'. A total of 32 subjects were evaluated on the level of attention and closeness between the speaker and the subjects. Descriptive results showed that participants of the visually co-located VMC condition performed better in attention level and had closer feelings toward the speaker than those in the control group, without a synchronized background setting. However, these results did not reach statistical significance. Thus, the effect of a synchronized background in VMC remains unclear. Future research should work on a long-term (i.e. months) experiment to see if there is any effect that requires time to reveal and furthermore, work on a clearer, less distracted research methodology.

Keywords: Video-mediated-conversation, attention, closeness, visual cue.

1. INTRODUCTION

1.1. Communication in the modern era

Sociability is what makes human beings survive the cruel nature environment. Our ancestors are known to gather together through social activities (e.g. grooming) to build connections among members (Gamble et al., 2014). Therefore, communication is necessary for the social need of humans. As time goes by, new methods are invented to pursue better ways to connect with each other.

In the 20th century, internet technologies have been introduced to the common world and have changed the concept of ordinary social life. Alongside the development of information and communication technology, we overcame the limitation of physical space. Making contact with others became much easier and faster. Mobile devices such as smartphones or tablets became part of our daily life and even part of our body. E-communication has led us to an era of perpetual contact, it allows us to be 'always-on' and 'never offline' (Katz & Aakhus, 2002).

1.2. The key role played by video-mediated communication (VMC)

E-communication is inevitable for contacting nowadays. Ever since video telephony was introduced to the public and became common for individuals, it has enabled us to virtually keep in touch with our family members or friends all over the world. With numerous technologies, we are capable of things we could not have accomplished before. Modern society is free from the physical limitations that prevent us from real-time communication. We do not need to be in the same place to start a conversation and neither do we need to travel all the way just to see each other's face.

Video-mediated communication (VMC) cannot be seen as a transit tool, which will no longer exist in the future or as just a substitute for FTF communication. The potential for VMC is big, as it can be applied in multiple fields; it offers individuals an alternative option for private contacts and offers a new approach for fields like business and education.

Telecommuting, a cross-country conference and cooperative work based on e-communication, are now part of the working routine in many organizations. In education, distance learning has always been an option for learning. The history of distance learning can even be traced back to the 18th century (Subrahmanyam & Ravichandran, 2013). Although due to the communication limitations during these times, it was not a popular choice. With video telephony, it can not only improve the quality of distance education but also reduce the cost significantly. In this way, distance learning could provide equal quality courses, in comparison to the traditional, FTF method.

Due to the evolving quality of online video communication and its capabilities of instant responses and the visual presences, VMC could be a promising and economical method for teaching, cooperative works and every task requiring remote communication. To provide a good VMC setting which could help users focus on the communication itself is not only important to the effectiveness in education but also to the utilization of other fields. In order to do so, the attention issue cannot be overlooked, it could be the key to increasing the concentration and improving the communication quality.

1.3. Two opposite viewpoints toward e-communication

There are always objections and concerns to new technologies, their capabilities of changing the status quo are powerful and also terrifying to some people. In fact, similar objections against 'writing' showed up in Plato's time thousands of years ago. Plato and Socrates both were against this 'manufactured product', which is inhuman and could weaken minds in their opinion (Miller & Sinanan, 2014, p. 4). The same arguments against e-communication can easily be seen nowadays as well. For example, people ask questions like "Why we don't talk to each other anymore? (John Locke, 1999)" or "Why do we expect more from the technology and less from each other? (Turkle, 2017)". These arguments reveal the concern of using e-communication as it only does harm to the social relationships. However, this is just one side of the story.

In Turkle's latest book: *Reclaiming Conversation: The Power of Talk in a Digital Age*, she (Turkle, 2016) states a relatively passive point of view towards the current progress of communication in the digital age. She mentioned two research findings showing that even an

insignificant, silent appearance of a smartphone can inhibit the development of intimacy and trust (Przybylski & Weinstein, 2013) and make people feel less satisfied in a FTF conversation (Misra et al., 2016).

On the other hand, Miller and Sinanan (2014) hold an optimistic view of the technology influence on relationships. They summed up their studies into a theory of attainment in their book, *Webcam*. Their studies suggest that people should take a rather neutral viewpoint on this topic and consider new technologies as an attainment; an equipment we could use on ourselves to acquire or extend some ability as well as to achieve tasks we could not have done before.

However, the arguments and discussion should not result into a binary opposition that either of e-communication harm our ability to be social, or they actually do us good and improve our ability as social animals. Even though Miller and Sinanan are rather optimistic about the future of people in the digital age and show more faith in the social relationship development, they also emphasize that it does not mean that the new attainments people acquire, would not cause any negative effect on themselves (Miller and Sinanan, 2014, pp. 1-23).

In fact, in one of Turkle's FTF interviews of a WhatsApp Group of 25 young people, Turkle found out that the group members were constantly distracted and always 'elsewhere' (Turkle, 2016, p. 35). It was hard for them to build up deep conversations because there were too many parallel activities going on both in reality and online. Their attention was constantly switching between one and another. As Turkle mentioned again and again in her book, the distracted mind and the disability of sympathizing with others are what hurt the conversation and emotional bonds within social contacts the most (Turkle, 2016).

Therefore, the issue of digital communication appears not to be 'how digital communication alienates people', but rather focusing on an individual's 'attention' level within conversations. The digital devices have cultivated new social behaviours in human, which lead to distracted minds.

1.4. Research design and methodology

Previous researches on VMC focused more on the emotional status of people doing VMC and the comparison of VMC, FTF and other e-communications. For example, the intimacy couples

felt during the video chat with their long-distance relationship partners (Neustaedter & Greenberg, 2012) and how people react differently in FTF and VMC (Shin, Jang & Bente, 2017). Miller and Sinanan (2014, p. 93) showed evidence that people do concern about the background that they are showing on the VMC, but there is no research that focuses on the link between the background factor and the attention level in VMC yet. This research aims to find if there is any connection between the background of VMC and the attention level of participants in the video call.

The hypothesis for this study is that a visually co-located video-mediated-communication (a synchronized background with the participants) can increase the attention level in VMC. An online experiment is built for this research with two different background settings in a simulated video call. Participants in the experimental group will see the speaker in the video call sitting in the same surrounding, with the same background as themselves. During the VMC, participants will have to react to audio stimuli (the word 'I') and the reaction time will be recorded for evaluating the attention level. After the VMC section, participants need to answer a knowledge test about the simulated video call and fill in the questionnaire for measuring their relatedness and pressure feelings during the previous section and define the closeness between themselves and the speaker.

This paper is divided into seven sections and appendices. Section 2 discusses in more detail about related works of VMC and attention; section 3 states the research question and hypothesis of this study; section 4 describes the methods used in testing the hypothesis; section 5 demonstrates the results and statistical analyses of the experiment; section 6 includes the discussions based on the experiment and the conclusion of this research.

2. RELATED WORKS

2.1. The framed nature of every communication method

Face-to-face (FTF) communication is often considered as the most natural, real and true way to make a contact with others. However, this is an illusion that we fail to break. Miller and Sinanan (2014) hold the same viewpoint as Goffman did a half-century ago (Goffman, 1959; Goffman, 1975), that there is no communication going on without being mediated. Every communication method has its own framework; no matter it is an e-communication or a FTF contact. Due to the fact that people are exposed to new technologies and multiple communication methods, the self-consciousness and the awareness about the frames of human interaction are increasing. The nature of every communication method is not more (or less) framed than in the old days, but we are definitely more aware of the frames within the communication. Since people are interacting through the 'windows' of their digital devices with each other, it is hard to not notice that personal communication is indeed happening within frames (Miller and Sinanan, 2014, p. 8). With the increasing consciousness of framed communication, the concerns about e-communication and how it might alienate people have risen as well.

Based on Darwin's evolution viewpoint, media naturalness theory argues that FTF communication is a survived result of natural selection and through the evolution progress. Human being have been 'engineered' to choose "a co-located and synchronous method, as well as through facial expressions, body language, and speech" (Kock, 2011) over other contact methods. To a certain extent, it does explain why most people fail to see the fact that FTF is mediated and not as natural as they thought. Following the arguments above, the next paragraph is going to compare the differences between FTF and e-communications (mainly of VMC) and provide evidence from previous researches.

2.2. The myth of FTF as a superior communication method than others

The most obvious limitation in computer-mediated-conversation (CMC) is that people cannot be physically co-presence with the subjects they communicate with. One of the advantages

of FTF communication is that it can provide the complete social context with nonverbal visual cues, such as body language, sounds and facial expression. These nonverbal cues are considered important and crucial components for reaching emotional intimacy in communication (Manstead, Lea & Goh, 2011, p. 147). Even though the media richness hypothesis argues that *“the face-to-face medium is the richest and most effective medium for reducing equivocality”* (Kock, 2011). Other research outcomes suggest that FTF is not necessarily the best or superior to the other communication methods.

Research findings have demonstrated that the methods (e.g. telephone, video call) other than FTF, instead of having a worse outcome, are actually able to achieve the same intimacy level as talking in person (Manstead, Lea & Goh, 2011). Face-to-face via the internet (i.e. VMC) can even produce a more emotionally intimate relationship between people than FTF with physical co-presence (Manstead, Lea & Goh, 2011, pp. 165-167). The research findings of Walther and Bazarova (2008, as cited in Walther, 2011, p. 30) proved that each type of medium can offer the same satisfaction if everyone in the small group (three- or four-person group) was interacting with the same medium. The emotional experiences of the groups using e-communications (e.g. video telephony, telephone, text-based CMC) showed *no less satisfied* than the *groups using FTF* communication.

In other words, FTF does not provide an absolutely more satisfied communication experience than any other media. The differences between mediums are not about the capability of offering participants a satisfied conversation, but the various operating manuals of each method. Participants need to dedicate efforts to learn how to apply their social skills into different communication methods. While using communication mediums other than FTF, they need to learn and be familiar with the ‘rules’ or ‘common sense’ of the medium. For instance, a smiley face emoji in texting can have completely opposite meaning within different contexts. Therefore, in order to deliver their message correctly and efficiently, people have to adapt to the ‘common sense’ of the media. In this way, users are able to transfer their thoughts into the form of the media, for instance, the text-based CMC users can *“adapt their multiple meanings into the single channel of language online.”* (Walther, 2011, p. 28).

Galagher and Kraut (1994) conducted an experiment of three different communication conditions on students: FTF, text-based CMC and text-based CMC with telephone as an extra method. The students had to use the media assigned to them to finish their team project. The

results showed that in the text-based CMC condition, those allowed to have phone connection with their partners already had more productive and enjoyable talks. This indicates the closer to a FTF communication condition, the more satisfied people could be. However, the quality of their outcome (i.e. the grades of their teamwork) in these three conditions did not varied from each other (Galagher & Kraut, 1994, p. 74). Another piece of evidence (Dennis & Kinney, 1998) indicated that there was no significant difference in the outcomes (i.e. the decision quality, consensus, satisfaction) of text-based CMC and VMC between the control and experimental groups. The richer media only showed its effect on supporting a faster decision-making progress (Dennis & Kinney, 1998, p. 269).

These findings indicate that even though subjects seemed to consider a richer media more satisfied, the quality of their outcomes from different communication methods (i.e. FTF, CMC, VMC) did not differ from each other. On the other hand, it suggested that increasing the media-richness (e.g. non-verbal visual cues) could help improving the satisfaction and intimacy level of the conversation.

2.3. Presences in VMC: physical and social dimensions

“Seeing is believing.” Despite the saying, this is not a true statement in general, it still suggests that visual cues can give us a certain base of building *authenticity*. As mentioned before, *nonverbal visual cues* are important for building emotional connection and intimacy, especially the facial expression. While making contact, the sense of presence is a necessary factor in building a communication. There are two aspects of presence (Manstead, Lea & Goh, 2011, p. 149). One is the *sense of physical presence*, which means being able to see and be seen by another. In other words, the *awareness* of another and the capacity to relate to them. The other one is the *sense of social presence*. That is, the *social meaning* of the context conversational partners can get from the communication.

According to Short, Williams and Christie (1976), social presence was first defined as “the salience of the other in a mediated communication and the consequent salience of their interpersonal interactions” (Short, Williams & Christie, 1976, p. 65). There are two components tightly linked to it: *intimacy* and *immediacy*. *Intimacy* (Argyle & Dean, 1965; Tu & Mclsaac, 2002) means the emotional closeness one feels. *Immediacy* (Wiener & Mehrabian,

1968) is used to reflect interaction intensity and it is a process to build up the desirable social outcomes.

If we take a closer look at VMC, it fulfils all factors in social presence. Besides, physical presence is not a necessary element for establishing a rapport (Manstead, Lea & Goh, 2011). Social presence is more crucial than physical presence because the physical presence alone does not necessarily build up a conversation.

2.4. A few evidence of VMC degrading the communication quality

Based on the arguments above, VMC is a communication method which would not degrade the conversation quality but is a reliable method helping us keep contacts with people from far away. However, there were several studies suggested otherwise.

While the research of Muhlfelder et al. (1999, as cited in Walther, 2011, p. 23) found out that interpersonal trust does not differ from the communication method used, and that video telephony is not significantly different from FTF. On the other hand, the comparative research of Storck and Sproull (1995) found out that in a classroom setting experiment, the participants using VMC, tend to undervalue other classmates more than the participants using FTF. This study suggested that under a teaching scenario, video telephony was unable to build the emotional bonding at the same level as FTF method and it could even degrade the quality of the communication.

Technical restraints could be a factor degrading the VMC communication as well. The audio and video system Storck and Sproull (1995) used in their setting two decades ago, were not synchronized perfectly. Previous studies showed that the transmission delay could be disruptive while establishing a communication. Although the mismatched audio and video signals did not have a significant negative influence on those who already knew each other before the experiment. The delaying signals could cause difficulties in conducting conversational task such as making a speech or giving up on the current talk. As a result, it reduced the sense of social presence and make participants felt less involved (Parkinson & Lea, 2011).

While comparing with FTF, video telephony under certain conditions (e.g. delayed signals and a class/group VMC setting) could degrade the experience of the communication or have a negative effect on the conversational partners in some cases. Nevertheless, most research findings still show that VMC could achieve the same emotional communication experience as FTF.

2.5. Attention: the basis of perception and interaction

Attention is the basis of human cognitive functioning for further perception and interaction with the world outside themselves. Without ‘paying attention’, people are not able to learn, reason, analyse and to do all the things which require cognition. In other words, attention allows us to interact with others consciously (Kindlon, 1998, p. 72). Therefore, attention measurement is included in this study as an indicator of psychological involvement in the communication.

There are multiple methodologies and experiments developed within the scope of attention, especially in the psychology field. All methods of measuring attention can be roughly divided into two genres: qualitative and quantitative measurements. Due to the fact that a behavioural and cognitive process such as attention is happening within individuals’ mind, it is always a challenge for researchers to find an appropriate approach in measuring it (Chaffee & Schleuder, 1986, p. 77). As a result, qualitative methods are limited to the *self-report* measurements, such as interview procedures or asking subjects to fill in the questionnaire from their own observation toward themselves; in the latter case, it can result in quantitative outcomes by grading their emotional feelings (Chaffee & Schleuder, 1986, p. 78).

Quantitative measurements are tests like simple reaction time test¹, the flanker test², the Ruff 2 & 7 selective attention test, behavioural observation (e.g. eye gaze direction and range, coding facial expression), psychophysiological measurement (e.g. measuring biological symptoms like blood pressure and brain waves), and so on (Kindlon, 1998; Chaffee & Schleuder, 1986; Baron, 2004; Jones et al., 2016).

Previous research usually combines both qualitative and quantitative methods in their study. Quantitative measurement for gathering numerical data and objective proofs, while self-report is used for checking the validity with outcome of the other measurements (Chaffee & Schleuder, 1986, p. 78). Therefore, in addition to the quantitative measure of a selective attention test, qualitative measures of the attention paid to the conversation and the mental status are also used in the experiment. In this study, we focused on the selective attention³ to see if the participants' attention was on the conversation or not. A part of the measures in the experiment was adapted from the framework of Ruff 2 & 7 selective attention test. Section '4. Methods' will discuss more in detail of the measurements in this study.

¹ Simple reaction time (SRT) is for measuring motor reaction time (RT) of responding to every stimulus presented in the test (Jones et al., 2016).

² The flanker test aims to test selective attention and the performance of the test reflects the participant's executive attention control. Participants respond to the central target accordingly in two conditions: with the flanking target on two sides the same as the central stimuli (congruent) or different (incongruent) from the central one (e.g. *HHHHH* vs. *HHSHH*) (Jones et al., 2016; Wilson, 2015).

³ According to Baron (2004), there are four types of attention: selective or focused attention, divided attention, sustained attention and alternating attention/mental shifting.

2.6. Conscious of location: 'background' factor in VMC

Recent research of using VMC in private life showed that people tend to choose a fixed place and stay there while having a video call, even if they were able to move around with their laptop (Kirk, Sellen & Cao, 2010). The moment people start a VMC at home, they literally open a *window* to their privacy for their conversational partners; it is a 'virtual invitation' for entering their private space (Miller & Sinanan, 2014, p. 93).

Despite the fact that video telephony is based on virtual connections, the actual location of conversational partners still matters to both parties. While having a video call, people become aware of their surroundings and are concerned about what their conversational partners can see on the screen. Furthermore, many users tend to 'manipulate' their own surroundings just to show their conversational partners how their place looks like, such as adjusting the recording angles of their webcams to show a cleaner corner of their room (Miller & Sinanan, 2014, pp. 92-95). Information about the speaker's *location* and what is *visible* there become crucial non-verbal visual factors in video conferencing. VMC is demonstrated on a two-dimensional screen (i.e. a video call window in a computer), thus the word 'background' in this study refers to the visual information of conversational partner's location.

To sum up, VMC as a communication method is not lower-class or less valued in most scenarios. In order to know the emotional effects that VMC background factor can have, not only the inner emotional status but also the attention level is taken into account.

3. Research Question

If the 'background' in video-mediated conversation (e.g. the place one is at.) is something that people are aware of during the video call, then theoretically speaking, modifying the background should be able to affect the experience people get from the VMC.

While thinking about a way that can make conversational partners feel like they are closer to each other, the 'synchronized' background idea came up; make the other conversational partner sitting in the same space as you. In other words, modifying the conversational partner's background into the subject's own surrounding in VMC, to see if this setting can rise subject's attention level of the conversation. This idea came to the research question:

"Does visually co-located video-mediated-communication draw more attention to the conversation?"

The hypothesis is that putting someone virtually in the same place as the subject can create stronger senses of presence and connection. As a result, enhancing the attention level on the talk. Which means, people are more engaged in the conversation while their conversational partner is in the same place as them.

4. Methods

4.1. Task

During the video-mediated conversation section, the test subjects' task was to indicate at which point the speaker says the target words (i.e. stimuli) by clicking on the react button on the screen. Afterwards, the subjects needed to fill in a survey of personal information and experiences about using VMC: a self-report and evaluation survey. In the survey section, participants would be tested on knowledge of the talk, including visual presence, contents, stimuli as well as their own subjective experiences during the VMC section.

Our experiment purpose is to see if the change in background affects the selective attention of test subjects. The selective attention includes two processes going at the same time: the inhibition of distractors as well as the focus of attention on stimuli (Wilson, 2015). Considering that the storytelling contains a lot of information already, there was only one specific word ('I') used as the stimulus in the experiment, while no other word was chosen for distractor. Under the audio content circumstance, every word is an information item popping out, therefore, the other words should be taken as distractors.

4.2. Procedures

The experiment consists of two parts, firstly, an attention test in VMC; secondly, the survey section. The experiment is conducted under the situation similar to a real-life video call. It means participants would go through the whole online procedures by themselves, in their residential place (i.e. home), with their own computers (e.g. laptops or desktop computers) rather than conducting it in a lab environment. There are two groups with different background setting in the experiment (*Table 1*), the *neutral* setting (control group) and the *synchronized* setting (experimental group).

	Control group	Experimental group
Background setting	Neutral setting (3 randomly assigned neutral backgrounds.)	The conversational partner's video call background is synchronized with (the same as) subject's own surroundings.

Table 1 Group background settings

4.2.1. Procedure flowchart

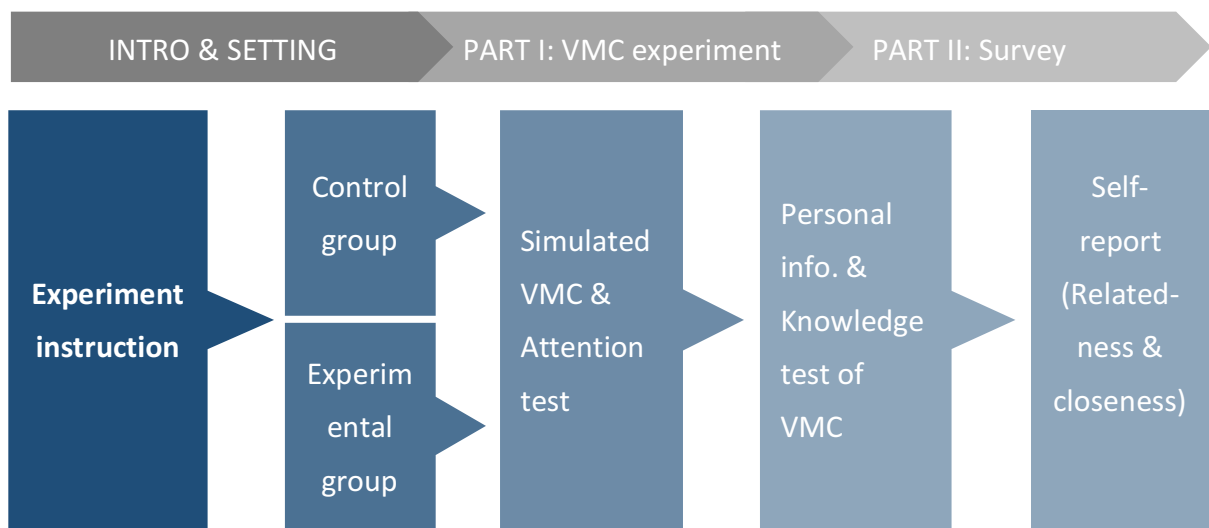


Figure 1 VMC experiment procedure flow chart

Figure 1 above illustrates the conducting structure of the experiment. There were three sections in the experiment. The first section was introduction and experiment instruction, in which participants were asked to read guidelines and check if their PCs and web browser meet the technical requirements of the coming simulated VMC test. The second was VMC trial part, in which participants watched a simulated video call while reacting to the stimuli and the timestamps data were recorded to the server. After the simulated VMC, participants were going to fill in the survey based on the knowledge and perceptions toward previous trial and their personal information as well as experience about VMC.

4.2.2. Participants and sampling

Considering video call is common and popularized to the general population, the experiment used snowball sampling as the sampling method. The participant recruitment was sent

through my social network and posted in several survey exchange communities. Participants were informed with basic experiment information along with a lottery chance of winning a 25-euro reward after completing the whole experiment and asked to visit the experiment website.

4.2.3. Introduction and setting of the experiment

The experiment was a one-person experience. After entering the experiment website, participants would see the instruction section explaining how the experiment would go and the task they were supposed to fulfil during the call. They were also informed that there is no audio and video recording during the call. After reading the rules and allowing the data access terms, each subject could click to start and join the experiment. *Figure 2* shows the settings of the control and experimental group. Participants would be randomly assigned to a group and following the instruction to set up the webcam. The smiley face symbols in *Figure 2* and *Figure 3* indicate the position of the subject's live camera in webcam access page and the simulated VMC page. If they were in the experimental group, they were asked⁴ to take pictures of their own background for the setting (see *Figure 2*).

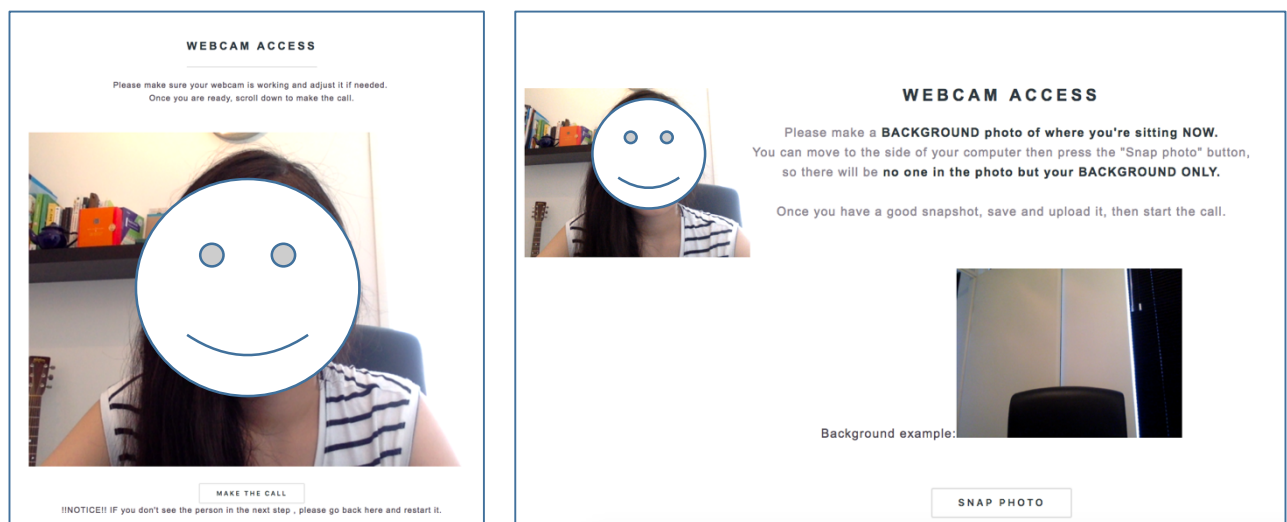


Figure 2 Control group (left) and experimental group (right) camera settings. The smiley faces in the figure represent participant's live camera position on the webpage.

⁴ Participants who were assigned to the experimental group could still choose to start the next VMC section without uploading their background photo in the webcam access page. In this case, they would be reassigned to the control group.

4.2.4. Part I: the VMC experiment

The first section is conducted in a simulated VMC environment for 2 minutes. It contains a storytelling part and a real-time attention examination. Participants were instructed to react to the target word (i.e. stimuli) by clicking the react button before joining the simulated video conference with either neutral setting or synchronized setting. When it started, the participant would see a woman sitting in front of her webcam and telling a story as if she was talking to the participants.

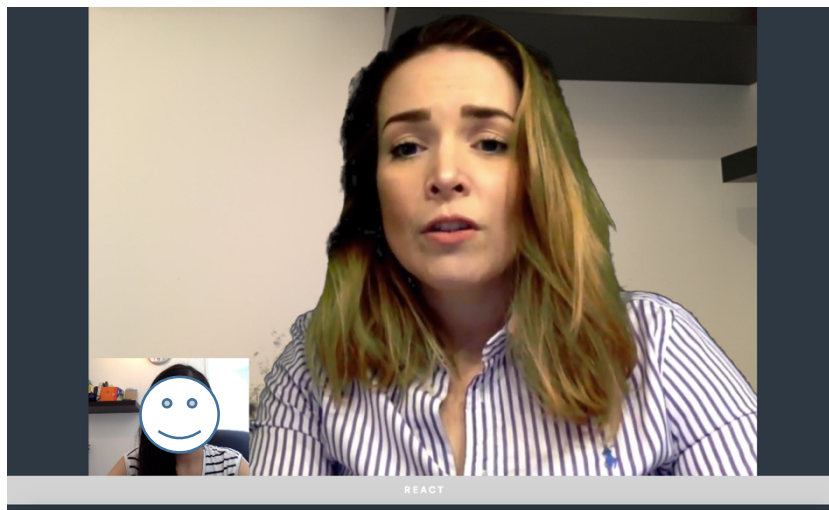


Figure 3 The simulated video call screenshot of the control group (non-synchronized background) with the react button underneath. The smiley face represents the position where the participant would be.

Figure 3 illustrates how the simulated VMC was like in the experiment for participants in the control group. The only difference in the call was whether test subjects' background was the same as the speaker's. The speaker had two neutral backgrounds for control group scenario use (*Figure 4*). The speaker is a 30-year-old Caucasian Dutch woman who speaks fluent English in the call. Meanwhile, the participant needed to listen to what she was going to say and press the react button when the stimulus word comes out while avoiding other distracting words.



Figure 4 The speaker's two neutral background settings for control group scenario (left: bg0.jpg; right: bg1.jpg).

The script used in the simulated VMC film is from Simon Corcoran, an ex-examiner of IELTS (The International English Language Testing System). It is a sample (band 9, highest score in IELTS) answer following the standards of IELTS speaking test part 2 (Corcoran, 2013) (see *Appendices 8.1*). It is a one-way and complete storytelling fulfilling the standards of one of the biggest English language test in the world⁵, moreover, its story structure⁶ serves the purpose of the experiment to have an appropriate, fluent and complete speaking segment with details for further understanding-of-the-conversation test (i.e. knowledge test) in the second section. The target word here is chosen as the first person 'I' of personal pronouns and the distractors are other personal pronouns or other words of the speaking in general, there are 14 'I's in the script.

⁵ IELTS introduction (n.d.), retrieved April 26, 2018, from <https://www.ielts.org/what-is-ielts/ielts-introduction>.

⁶ The sample script was meant for the IELTS speaking test which means it meets the standard of IELTS scoring system. Therefore, it is a qualified script to use in the speaking. (How IELTS is scored. (n.d.). Retrieved April 26, 2018, from <https://www.ielts.org/about-the-test/how-ielts-is-scored>)

4.2.5. Part II: the survey

Part two is to fill in a survey and answer relevant questions about the previous section to the participants' best memory. This survey consists of a knowledge test, self-report of subject impression (IMI measure) and closeness (IOS scale), personal information such as age, English self-assessment⁷ and VMC-using experience. The knowledge test is a multiple choices test, in which subjects have to choose whether the description of the simulated VMC is correct or not. As for self-report part, it requires subjects to answer several questions of the closeness and relationship feelings toward the speaker.

4.3. Experiment Environment

Due to the research design, the whole experiment was built on a website for the participants to do it under actual VMC condition: a remote video call. The website⁸ consists of certain HTML5 features, such as local storage for recording each click time of the participant⁹, webcam access, photo capture¹⁰ for synchronized background use and so on. After gathering the data needed and saving them in the web server, the program would automatically clear all the data saved on the client side, for the security and privacy reasons.

⁷ Self-assessment options are extracted from the self-assessment grids of Common European Framework of Reference for Languages (CEFR). Only the English listening descriptions from B2 and C2 (highest) level were used in this survey. (Self-assessment Grids (CEFR). (n.d.). Retrieved May, 2018, from <https://www.coe.int/en/web/portfolio/self-assessment-grid>)

⁸ The website (<https://leepinhhsien.nl>) is made with the theme "Spectral" from HTML5 UP (<https://html5up.net/>, May, 2018).

⁹ This function related to local storage is adapted from an online code sample which creates items based on text input (Tania Rascia, Dec. 18, 2017) instead of timestamps in VMC experiment website.

¹⁰ This function related to camera and saving webcam screenshot is adapted from an online code sample (Walsh, Sep. 23, 2016).

4.4. Measures

There are four measures in this experiment, the first one is inspired by The *Ruff 2 & 7 Selective Attention Test*, the second is knowledge test score (KTS) of the simulated video call, the others are following the measures of *Experiment I* in the research of Przybylski and Weinstein (2013).

4.4.1. Attention level in real time: reaction time

In order to test the selective attention with higher construct validity and external validity, meanwhile using the method suiting our context and goal the most, we decided to conduct an experiment inspired by *Ruff 2 & 7 Selective Attention Test* after reviewing all the existing means by now (Wilson, 2015).

The *Ruff 2 & 7 Selective Attention Test* aims to evaluate the visual selective and sustained attention by asking subjects to cross out as many '2' and '7' numbers as possible where they are printed on the page alongside with letters in 15 to 20 seconds (Ruff et al., 1992). While conducting our experiment, we have to transfer the core concept into video-mediated-conversation testing environment, which means subjects need to 'cross out' the target 'word' (i.e. stimuli) in an ongoing VMC as those who cross out '2' and '7' on the paper in the *Ruff 2 & 7 Selective Attention Test*, while having other distractors in the conversation.

4.4.1.1. Scoring system of reaction time

In order to score the test results, there are three scores taken into account: the *reaction time score (RTSUM)*, *accuracy percentage (RTAC)*, and *distracted rate (RTDISTRACT)*.

RT (seconds after stimulus)	0.0 - 0.6	0.7 – 1.3	1.4 – 2.0	> 2.0 – 5.0	React to non-stimulus*
Result	Correct (the first time)			Error	Error
Score	7	4	1	0	1

Table 2 Reaction time scoring system.

(1) Reaction time (RT): sum of reaction time scores from each correct hit (see *Table 2*).

Subjects needed to react (i.e. clicking on the button) in a 2 second time range after noticing the stimuli. Only *the first hit counts for a hit and scores* within the range¹¹. Therefore, the highest possible score is 98 ($14 \times 7 = 98$). Same goes for counting error, if the errors are made within 2.0-5.0 seconds range, only the first one counts as an error hit. If the Reaction time is over 5 seconds comparing with the previous stimulus, then it is counted as an error.

Although there is one gap between 12th and 13th stimulus is 1.2 seconds, the pre-test results show that most subjects react to the stimulus in 2 seconds range. Therefore, if it used 1.2 seconds for time range, the scores from subject would be too low to indicate the actual situation. Considering the other time gaps are equal or more than 2 seconds, using 2 seconds range for scoring is more suitable for the test score. As for the overlap 12th and 13th gap, there are 4 situations for scoring:

- a. If there is only ONE click within this 3.2 seconds range:
 - It will be counted as the 12th score while it is within 2 seconds after 12th.
 - It will be counted as the 13th while it is within 2 seconds after 13th, but already passing the 12th range.
- b. If there are MORE than one click within this 3.2 seconds range (see *Figure 5*):

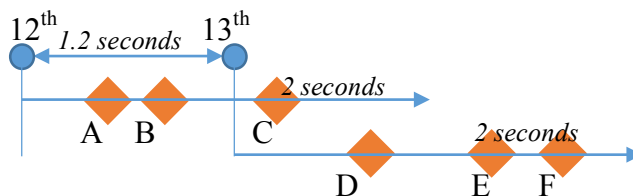


Figure 5 Scoring system for 12th and 13th stimulus (A-F are possible clicking positions.).

- If they both fall into the 12th OR the 13th time range (e.g. A&B or E&F in *Figure 5*) only, do not overlap with other time range, then only the first one counts (as the general rule).

¹¹ For example, if there are three hits falling respectively into 7-score range, 4-score range and 1-score range, the person only gets the highest score (i.e. the closest hit) of his hits within 2 seconds range.

- If they fall into 12th AND 13th, which means each range gets at least one react click (e.g. C&D, B&D, D&E...etc.), then both count for two stimuli (e.g. if they are C&D, C counts for 12th and D for 13th). If there are more, then the scoring follows the general rule that only the first one count for one stimulus.

(2) Accuracy percentage (%): $\{(total\ correct\ hits) / stimuli\ amount\} * 100$

The first hit corresponds to a single stimulus within 2 second time range count as one correct hit, the other times within range does not count as valid hits.

(3) Distracted percentage (%): $(total\ error\ hits) / stimuli\ amount * 100$

If subjects react to non-stimuli (distractor), no matter the mistakes are made consciously or unconsciously, they are considered as error hits. That is, hits outside the reaction time range (within 2 seconds after a stimulus) are all considered as error hits.

4.4.2. Attention paid to the conversation: knowledge test

In the survey, subjects answered 15 multi-choice questions about the video call they just had. Questions include both visual (4 QNS) and content (11 QNS) components, such as “What did she wear in the video call?” and “In her story, who was the person she had a conversation with?”. The score from this part of the survey is taken as knowledge test score (KTS), which indicates the attention paid to the conversation.

4.4.3. Relatedness and pressure/tension score

There are seven dimensions in the Intrinsic Motivation Inventory (IMI) questionnaire. In this study, the questionnaire used here is adapted from the ‘Relatedness’ and ‘Pressure/Tension’ dimensions only. As McAuley, Duncan and Tammen (1989, p. 49) notified in their study, the complete scales were rarely used and results showed that neither inclusion nor exclusion of the dimensions caused influences on the other chosen dimension(s).

IMI measure has been broadly used in testing subjects’ inner motivation status in voluminous study fields such as sports, video game, business organization and mental health disorder (McAuley, Duncan & Tammen, 1989; Ryan, 1982; Ryan, Rigby & Przybylski, 2006; Tremblay et al., 2009; Choi, Mogami & Medalia, 2010). It is a Likert-scale measure based on the agreement

degree from 1 (Not at all) to 7 (Very true), including 8 items with descriptions¹² such as “I’d like a chance to interact with this person more often.”, “I felt really distant to this person.”, “I don’t feel like I could really trust this person.” etc.

4.4.4. Closeness

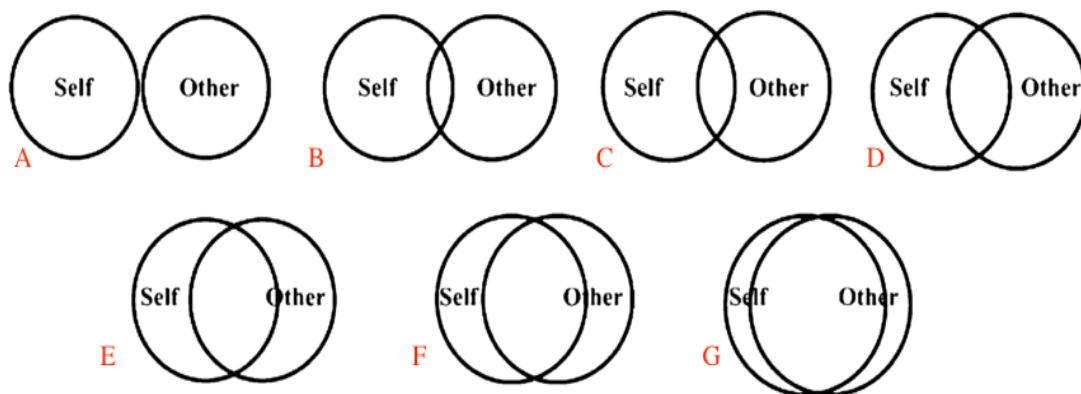


Figure 6 The Inclusion of Other in the Self (IOS) Scale.

Besides evaluating the attention level, the level of closeness during the VMC is also included in measurements. Previous studies have accepted the notion of closeness as *an overlap of selves* which inspiring the *Inclusion of Other in the Self Scale* (Aron et al., 1992) to come up. IOS scale (Figure 6) was used for measuring the closeness between the participant and the speaker in simulated VMC film. IOS scale aims to clarify one’s *sense* of being interconnected with each other, test subjects are instructed to select one pattern out of seven increasingly overlapping circle pairs indicating the involvement level of themselves and the speaker.

¹² Despite the fact that IMI measure is used everywhere in researches, the original questionnaire is lack of it is official resource to gather, except for a document we found on the website of Self-Determination Theory (Intrinsic Motivation Inventory (IMI), retrieved April 26, 2018, from <http://selfdeterminationtheory.org/intrinsic-motivation-inventory/>) and in the book, *Cybernetic approach to project management* (Ryan, Rigby & Przybylski, 2006, p. 381)

5. RESULTS

As *Figure 7* illustrates below, there were 32 test subjects in this experiment: 19 people in the control group and the other 13 people in the experimental group. Participants included 24 female and 7 male subjects, while one subject preferred not to reveal his/her gender. Test subjects were from 18-54 years old, most of them were from 24-34 age group and 27 test subjects conducted the experiment at home (other options are at school or office.), only one of 32 subjects was interrupted in the experiment for 1-2 times.

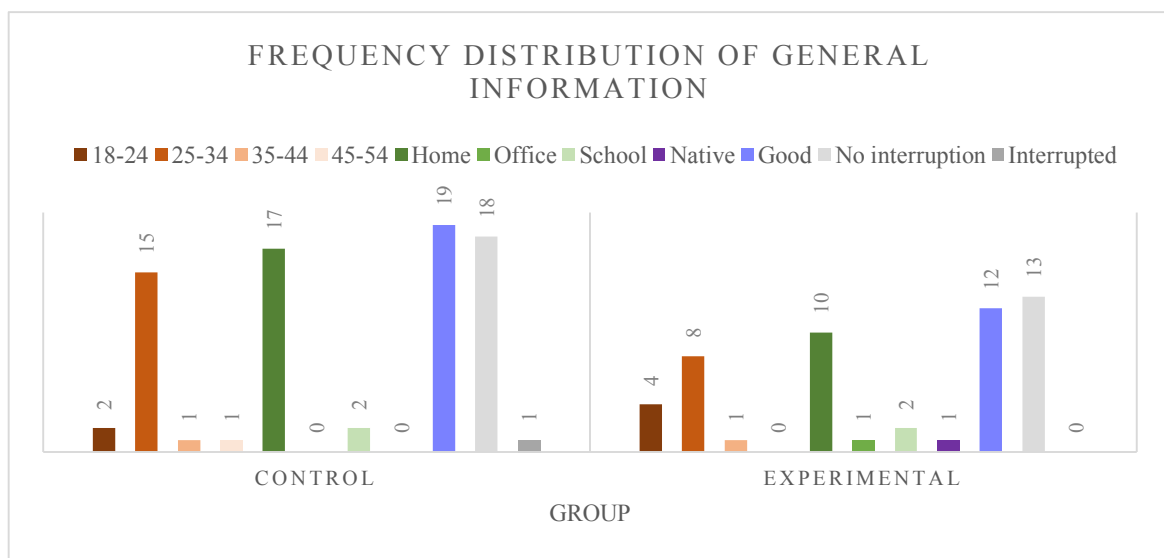


Figure 7 Frequency distribution table of general information includes gender, place, English level, was interrupted or not.

In order to prevent the effect causing by knowing each other, testing subjects were asked if they have ever seen the speaker in the simulated film before the test. If their answers to it were yes, those data would be invalid and discarded. In the end of the experiment, all 32 subjects' data we received were valid data¹³. All the variable abbreviations used in analysis are demonstrated in *Table 3*.

¹³ Test subjects might have doubt for whether they saw her before or not. If they choose "Maybe, but I don't really know her." as their answer, they'll still be taken as valid subjects.

Abbreviation	Score meaning & range	Explanation
VMCFREQ	'Rarely' to 'Constantly' (0-7 points)	VMC-using frequency score.
PCFREQ		Frequency score of using PC for VMC.
PFREQ		Frequency score of using mobile phone for VMC.
TABFREQ		Frequency score of using tablet for VMC.
KTS	0-15 points	Knowledge test of the simulated video call content.
PRS	0-35 points	Pressure/Tension score.
RELATEDS	0-56 points	Relatedness score
IOS	Figure 'A' to 'G' (0-7 points)	The inclusion of other in the self-scale (IOS) scores.
RTSUM	0-98 points	Reaction time test score.
RTAC	0-100%	Accuracy percentage in reaction time test.
RTDISTRACT		Error percentage in reaction time test.

Table 3 Abbreviations and the scores of variables used in the analysis

5.1. VMC-using experience

The variables in *Table 4* are subjects' VMC-using frequency in general (VMCFREQ) and respectively on three types of devices in the following order: PC, smartphone, tablet. As the chart shows below, p values are all above 0.05, that is, there's no difference in user behaviour of VMC between control and experimental groups. Based on the Shapiro-Wilk normality test (see *Table 5*), most of the data are normally distributed except for the 'using phone as VMC device' variable ($p < 0.001$) in control group.

Independent Samples T-Test

	t	df	p
VMCFREQ	-0.050	30.000	0.961
PCFREQ	0.685	30.000	0.498
PFREQ	-0.404	30.000	0.689
TABFREQ	1.080	30.000	0.289

Note. Student's t-test.

Table 4 Student's t-test of VMC-using experience variables in control and experimental groups.

Test of Normality (Shapiro-Wilk)

		W	p
VMCFREQ	Control	0.814	0.002
	Experimental	0.886	0.087
PCFREQ	Control	0.866	0.012
	Experimental	0.917	0.229
PFREQ	Control	0.795	< .001
	Experimental	0.844	0.024
TABFREQ	Control	0.851	0.007
	Experimental	0.816	0.011

Note. Significant results suggest a deviation from normality.

Table 5 Shapiro-Wilk normality test of VMC-using experience variables in control and experimental groups.

5.2. Variables results

Descriptive Statistics

	KTS		PRS		RELATEDS		IOS		RTAC		RTDISTRACT	
	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental	Control	Experimental
Valid	19	13	19	13	19	13	19	13	19	13	19	13
Missing	0	0	0	0	0	0	0	0	0	0	0	0
Mean	9.789	9.846	16.789	17.308	34.316	35.231	1.053	1.231	56.768	64.823	50.758	42.854
Std. Deviation	3.599	2.940	5.692	7.878	6.307	8.084	0.970	0.927	19.559	16.612	17.469	13.354
Minimum	2.000	4.000	8.000	5.000	24.000	18.000	0.000	0.000	21.400	35.700	28.600	21.400
Maximum	15.000	14.000	28.000	29.000	46.000	48.000	3.000	2.000	85.700	100.000	92.900	71.400

Table 6 Descriptive Statistics of multiple variables. (See Table 3 for the explanations of the abbreviations.)

Table 6 and 7 include reaction time related variables, which show the attention level in real time, knowledge test score (KTS) indicating the attention level in conversation after the video call, relatedness and pressure/tension scores from IMI measure and IOS scale. Hypothetically speaking, KTS, reaction time (RT) related variables are dependent variables in this study, PRS, RELATEDS and IOS are intervening variables.

Independent Samples T-Test

	t	df	p
KTS	-0.047	30.000	0.963
RELATEDS	-0.359	30.000	0.722
PRS	-0.216	30.000	0.830
IOS	-0.519	30.000	0.607
RTSUM	-1.122	30.000	0.271
RTAC	-1.214	30.000	0.234
RTDISTRACT	1.377	30.000	0.179

Note. Student's t-test.

Table 7 Independent samples t-test of multiple variables. (See Table 3 for the explanation of the abbreviations.)

Test of Normality (Shapiro-Wilk)

		W	p
KTS	Control	0.952	0.420
	Experimental	0.956	0.698
RELATEDS	Control	0.941	0.274
	Experimental	0.948	0.574
PRS	Control	0.952	0.423
	Experimental	0.936	0.406
IOS	Control	0.848	0.006
	Experimental	0.720	< .001
RTSUM	Control	0.974	0.857
	Experimental	0.842	0.023
RTAC	Control	0.947	0.357
	Experimental	0.952	0.631
RTDISTRACT	Control	0.937	0.232
	Experimental	0.948	0.574

Note. Significant results suggest a deviation from normality.

Table 8 The Shapiro-Wilk normality test of multiple variables in two groups. (See Table 3 for the explanation of the abbreviations.)

The normality test results (Table 8) show that variables in both tables are normally distributed, however, the p-values from the Shapiro-Wilk normality test suggest that there is no significant difference in any of the variables between control and experimental groups,

which also means that PRS, RELATEDS and IOS have no intervening effect on dependent variables.

Pearson correlation analysis (*Table 9*) indicates several variables have correlations between each other and these sets of variables with correlations are listed in *Table 10*. For the obvious logical reason, the higher the accuracy participant had, the better the total score they earned. On the contrary, the more mistakes they made in the attention test, the worse their accuracy and total scores were. There is a large positive effect on the correlation between RTAC and RTSUM and large negative correlations of these two sets, RTDISTRACT and RTSUM, RTDISTRACT& RTAC.

		KTS	PRS	RELATEDS	IOS	RTSUM	RTAC	RTDISTRACT
KTS	Pearson's r	—						
	p-value	—						
PRS	Pearson's r	0.058	—					
	p-value	0.751	—					
RELATEDS	Pearson's r	0.045	-0.304	—				
	p-value	0.806	0.090	—				
IOS	Pearson's r	0.268	-0.162	0.045	—			
	p-value	0.139	0.375	0.805	—			
RTSUM	Pearson's r	-0.020	0.227	0.197	-0.076	—		
	p-value	0.911	0.213	0.280	0.681	—		
RTAC	Pearson's r	-0.006	-0.011	0.246	-0.008	0.871	—	
	p-value	0.976	0.951	0.175	0.965	< .001	—	
RTDISTRACT	Pearson's r	-0.177	0.115	-0.226	-0.054	-0.692	-0.857	—
	p-value	0.333	0.531	0.213	0.768	< .001	< .001	—

Table 9 Pearson correlation matrix of multiple variables

(See Table 3 for the explanations of the abbreviations.)

Pearson's r value indicates a small positive correlation between the closeness participants felt and their attention paid to the conversation. That is, participants who included the speaker more in themselves (i.e. having a closer emotional feeling of the speaker) performed slightly better in knowledge test. The same result goes to the pressure score and two attention test related variables, RTSUM and RTDISTRACT. Participants who had higher pressure during the VMC section had slightly better reaction time scores, however, they made more mistakes in the reaction task as well. The other sets with small positive correlation are RELATEDS and two attention test variables (RTSUM and RTAC). The results suggest that participants who had

tighter relatedness feelings of the speaker performed slightly better in accuracy and the total score of the reaction time test.

On the other hand, there are six sets having a negative correlation. The sets with a small negative correlation are the closeness and the pressure score, the knowledge test score and the distracted percentage, the relatedness score and the distracted percentage. That is, participants with higher pressure emotional status included the speaker less in themselves. Participants making more mistakes in the attention test paid a bit less attention to the content of the VMC section and they had slightly lower relatedness feelings toward the speaker.

However, the p-value results of Pearson's correlations shown in *Table 9* suggest that even if there is any weak correlation between variables, there is no significant correlation between most variables. Therefore, we cannot say that the correlations actually exist in those sets, except for these three sets of dependent variables: RTAC& RTSUM ($p<0.001$), RTDISTRACT & RTAC ($p<0.001$) and RTDISTRACT& RTSUM ($p<0.001$).

Effect size	Small	Medium	Large
Positive effect	IOS& KTS ($r= 0.268$); RTSUM& PRS ($r= 0.227$); RTDISTRACT& PRS ($r= 0.115$); RTSUM& RELATEDS ($r= 0.197$); RTAC& RELATEDS ($r= 0.246$)	None	RTAC& RTSUM ($r= 0.871$)
Negative effect	IOS& PRS ($r= -0.162$); RTDISTRACT& KTS ($r= -0.177$); RTDISTRACT& RELATEDS ($r= -0.226$);	RELATEDS& PRS ($r= -0.304$)	RTDISTRACT& RTSUM ($r= -0.692$); RTDISTRACT& RTAC ($r= -0.857$)

Table 10 Sets of variables with correlation

5.2.1. Attention

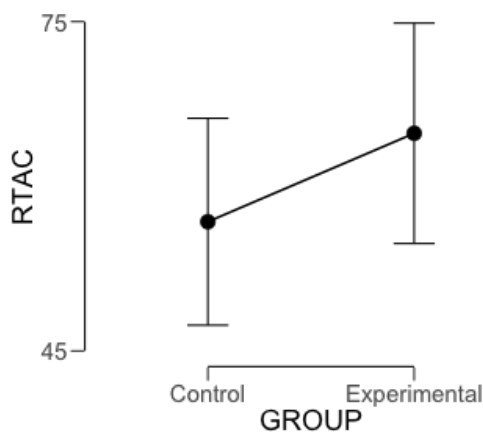


Figure 8 Reaction time accuracy descriptive plot of control and experimental groups. (RTAC: C=56.768, E= 64.823) (Within 95% confidence interval and the comparison of mean in two groups.)

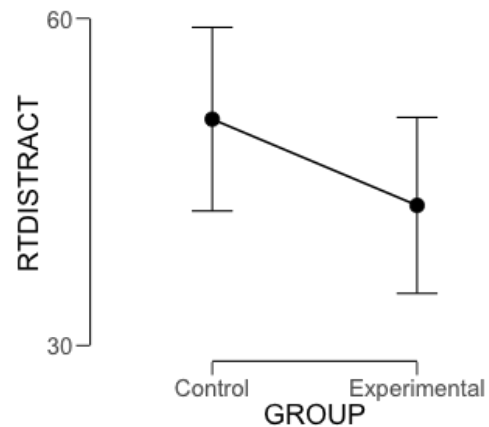


Figure 9 Reaction time distraction percentage descriptive plot of control and experimental groups. (RTDISTR: C=56.768, E= 64.823) (Within 95% confidence interval and the comparison of mean in two groups.)

Attention test related variables indicate the attention level in real time (Figure 8, 9), while the knowledge test score (KTS) indicates the attention paid to the conversation (Figure 10). Figure 8 shows that the experimental group had higher accuracy percentage mean than the control group (RTAC: C=56.768, E= 64.823), which means a visually co-located VMC condition helped participants to have a higher attention level during the call. Figure 9 indicates that participants in the experimental group made fewer mistakes (RTDISTR: C=56.768, E= 64.823), which shows that participants were more concentrated in the conversation when their conversational partner's background was the same as themselves.

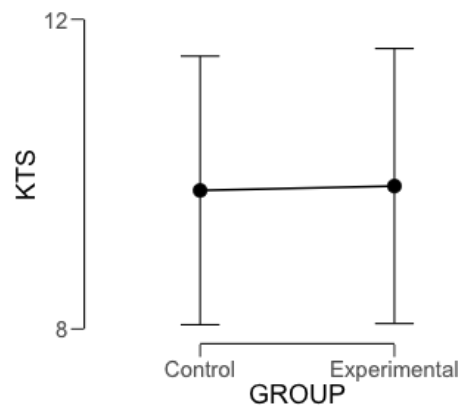


Figure 10 Knowledge test score (KTS) descriptive plot of mean in the two groups.

(KTS Mean: C= 9.7889, E= 9.846)

(Vertical bars indicate 95% confidence interval)

As for distracted percentage and KTS, *Figure 10* (see *Table 6* for the exact data) shows that the experimental group had slightly better score mean in knowledge test than the control group (KTS Mean: C= 9.7889, E= 9.846), which indicates participants in a synchronized background VMC condition paid more attention to the conversation than those who were not.

However, based on the p-value of Student's test in *Table 7*, the outcomes of all three variables (i.e. RTAC (p= 0.234), RTDISTRACT (p= 0.179) and KTS (p= 0.963)) did not reach statistical significance, thus we could not say the differences of RTAC, RTDISTRACT and KTS in two groups actually exist between the two groups.

5.2.2. Relatedness and pressure/tension score

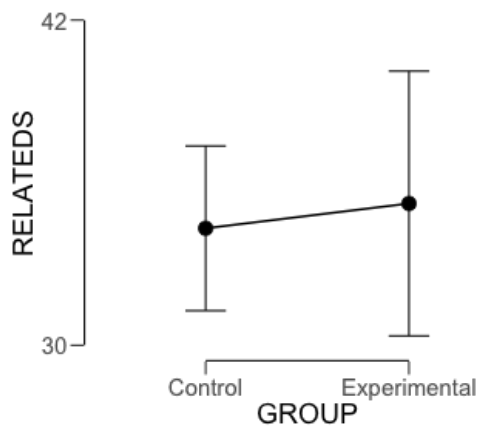


Figure 11 Relatedness (RELATEDS) descriptive plot of mean in the two groups. (RELATEDS Mean: C= 34.316, E= 35.231)
(Vertical bars indicate 95% confidence interval)

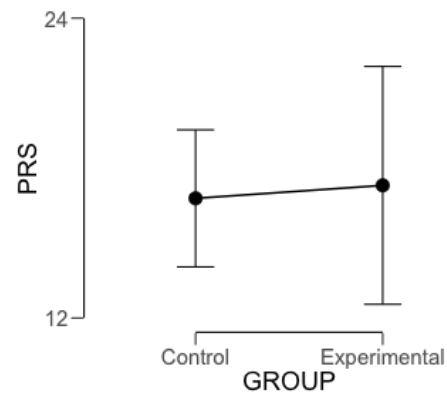


Figure 12 Pressure/tension score (PRS) descriptive plot of mean in the two groups.
(PRS Mean: C= 16.789, E= 17.308)
(Vertical bars indicate 95% confidence interval)

Similar results show in relatedness score (Figure 11) and pressure/tension score (Figure 12). The experimental group had a slightly higher mean (see Table 6 for the exact data) of both variables (PRS Mean: C= 16.789, E= 17.308; RELATEDS Mean: C= 34.316, E= 35.231). That is, a visually co-located VMC condition made participants related with the speaker a bit more than a normal video call setting, but they also felt tenser during the call. Nevertheless, according to the p-values of the RELATEDS ($p=0.722$) and PRS ($p=0.830$) in the Student's t-test (Table 7), the differences in the two groups were not statistically significant. Therefore, we cannot say these difference actually existed in them.

5.2.3. Closeness

Frequencies for IOS			
GROUP	IOS	Frequency	Percent
Control	A	6	31.579
	B	8	42.105
	C	3	15.789
	D	2	10.526
	Missing	0	0.000
	Total	19	100.000
Experimental	A	4	30.769
	B	2	15.385
	C	7	53.846
	D	0	0.000
	Missing	0	0.000
	Total	13	100.000

Table 11 Frequencies for IOS scale (see Fig. 6) in control and experimental groups.

(Data from E-G are empty therefore they are omitted in the table.)

The closeness score is based on the IOS scale, which contains seven circle figures describing the inclusion degree of others in self. From the most distant 'A' without any overlapped part to the closest 'G' that the two circles almost overlap each other. As *Table 11* shows above, over 50 % participants (7 out of 13) in the experimental group chose 'C', on the other hand, most people in the control group chose 'B' or the lowest 'A'. It suggests that participants in a synchronized setting tend to define the relationship between the speaker and themselves as a closer relationship than those who did not have the conversational partner being the same place as themselves.

However, as the Student's t-test results show in *Table 7*, the p-value of IOS is 0.607, which means the difference between the control and experimental groups did not reach statistical significance, thus we cannot confirm the difference here exist.

6. DISCUSSION AND CONCLUSION

6.1. Discussion

Despite the fact that the experimental group had better performance in the attention test (e.g. better accuracy and lower distracted percentage in the reaction time test and higher grades in the knowledge test), and closer feelings toward the speaker (e.g. higher relatedness score and a closer IOS figure was chosen) than the control group, these results did not reach statistical significance. Therefore, whether the background was synchronized or not, it did not truly have influence on the attention level in real time nor the attention paid to the conversation.

However, it does not mean that the synchronized background in VMC has no effect on people. There were some difficulties in the online experiment while conducting the trial and these problems could be the reasons why there was no significant outcome generated from the two groups.

The feedback and background data from the participants indicated that several problems occurred during the experiment, such as background selecting issues and the distraction from the reaction task. First of all, for participants, uploading their own background was a barrier that they needed to cross already. The uploaded background photos of the experimental group showed that some of the participants were not fond of the idea of (unwillingly) exposing themselves in front of the camera¹⁴. Even if the test subjects agreed to participate in the experiment, they were still very aware of the webcam and afraid of being recording by any means. This could be the reason causing the test subjects to choose a relatively 'safe' background for the experiment.

¹⁴ The landing page already indicated that there will be no recording in any audio and video format, but a few subjects still picked plain backgrounds to join the experiment (e.g. white wall). Some people even uploaded unqualified photos (e.g. cartoon picture, black wallpaper.) therefore, their data were not valid for the experiment thus discarded.

The results indicated that participants in the experimental group felt more stressed (see *Figure 12*, p. 32) than those in the control group. Even though the differences in these two groups were not significant, it could still suggest that the good performance in the attention test and the knowledge test had some help from the tense mental state. On the other hand, a few subjects thought that the reaction test in the VMC section distracted them from focusing on the call, which was totally the opposite of the initial intention.

The simulated video call was specifically made for this study. A lot of efforts were investigated into filming the simulated video call, making it look more natural and real. However, due to the design of the experiment, the VMC section could not be as spontaneous and flexible as a real video call. Besides the gap between a simulated VMC and a real VMC, the reaction task of the VMC section could have decentralise participants' attention as well. Instead of paying full attention to the talk itself, part of their attention could have gone to the task in the VMC section. Despite that the knowledge test was designed to avoid this exact problem, it may not be sufficient to prevent this distraction effect from happening.

6.2. Conclusion

Interacting with others is a necessary part of living, as a result, knowing the pros and cons of the method we use and exploring the possibility of it could actually help with building a better communicating experience for the modern and changing society.

In order to understand the factor and possible effect work behind VMC, this research looked into VMC from a different angle and approach: using an website to conduct a simulated VMC experiment for attention test. The descriptive statistical data showed that participants in a visually co-located VMC condition were more focused during the call and recalled more details of the call. They also felt emotionally closer to the speaker while listening to her in the simulated video call as well as perceived themselves had a more overlapped self and the other relationship with the speaker. However, the p-values of Student's t-test did not reach the statistical significance in all the outcomes mentioned above, therefore, the differences did not actually exist either.

Despite the fact that the experiment results failed to support the hypothesis that synchronised background gives a more immersive experience, hence bring you closer to the person you talk to. We believe it does not mean that a synchronised background has no effect on the closeness and the engagement in the video call. The known difficulties in examining and evaluating human's inside feelings (i.e. engagement, closeness and relatedness) make it tough to quantize and measure precisely for the research use. Therefore, in addition to quantitative and short-term approach, future studies could conduct a long-term experiment (e.g. 3 months) of a larger sample size and qualitative research approaches such as interviews. To see if the synchronised background needs time to exert its influence on people in the communication and understanding the influences in details by interviewing the participants.

From the trend of information and communication technology developments, we can see that a better, closer to real FTF contact is the ultimate goal for remote communication. Therefore, no matter how the information and communication technology evolves or the form of VMC changes, the core concept of VMC will not fade, but will only expand its coverage and be applied to more uses. Video-mediated conversation as a relatively new communication method still leaves a lot for us to dig into.

7. REFERENCES

- [1] Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry*, 28(3), 289-304.
- [2] Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, 63, 596–612.
- [3] Baron, I. (2004). *Neuropsychological Evaluation of the Child* (1st ed.). Oxford University Press.
- [4] Chaffee, S., & Schleuder, J. (1986). Measurement and Effects of Attention to Media News. *Human Communication Research*, 13(1), 76-107.
- [5] Choi, J., Mogami, T., & Medalia, A. (2010). Intrinsic Motivation Inventory: An Adapted Measure for Schizophrenia Research. *Schizophrenia Bulletin*, 36(5), 966-976.
- [6] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). *Hillsdale, NJ: Lawrence Erlbaum Associates*.
- [7] Corcoran, S. (2013, December 6). IELTS Speaking Part 2: Describe a conversation. Retrieved April/May, 2018, from <http://ielts-simon.com/ielts-help-and-english-pr/2013/12/ielts-speaking-part-2-describe-a-conversation.html>
- [8] Dennis, A. R., & Kinney, S. T. (1998). Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality. *Information systems research*, 9(3), 256-274.
- [9] Gallagher, J. and Kraut, R. E. (1994). Computer-mediated communication for intellectual teamwork: an experiment in group writing. *Information Systems Research*, 5, 110–138.
- [10] Gamble, Gowlett, Dunbar, Gowlett, John A.J, & Dunbar, Robin. (2014). *Thinking big: How the evolution of social life shaped the human mind*. London: Thames & Hudson.
- [11] Goffman, E. (1959). *The presentation of self in everyday life* (New ed.). ed., Doubleday Anchors Books ; A 174). Garden City, N.Y.: Doubleday & Company.
- [12] Goffman, E. (1975). *Frame analysis : An essay on the organization of experience* (Peregrine books). Harmondsworth [etc.]: Penguin.
- [13] How IELTS is scored. (n.d.). Retrieved April 26, 2018, from <https://www.ielts.org/about-the-test/how-ielts-is-scored>
- [14] HTML5 UP. (n.d.). Retrieved April, 2018, from <https://html5up.net/>
- [15] IELTS introduction. (n.d.). Retrieved April 26, 2018, from <https://www.ielts.org/what-is-ielts/ielts-introduction>
- [16] Intrinsic Motivation Inventory (IMI). (n.d.). Retrieved April 26, 2018, from <http://selfdeterminationtheory.org/intrinsic-motivation-inventory/>
- [17] John Locke (1999). *Why We Don't Talk To Each Other Anymore: The De-Voicing of Society*. Simon and Schuster.

- [18] Jones, S. A., Butler, B. C., Kintzel, F., Johnson, A., Klein, R. M., & Eskes, G. A. (2016). Measuring the performance of attention networks with the Dalhousie computerized attention battery (DalCAB): Methodology and reliability in healthy adults. *Frontiers in psychology*, 7, 823.
- [19] Kappas, A., & Krämer, N. (2011). *Face-to-Face Communication over the Internet: Emotions in a Web of Culture, Language, and Technology*. Cambridge: Cambridge University Press.
- [20] Katz, J., & Aakhus, M. (2002). *Perpetual Contact: Mobile Communication, Private Talk, Public Performance*. Cambridge: Cambridge University Press.
- [21] Kindlon, D. (1998). The Measurement of Attention. *Child Psychology and Psychiatry Review*, 3(2), 72-78.
- [22] Kirk, D. S., Sellen, A., & Cao, X. (2010, February). Home video communication: mediating 'closeness'. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 135-144). ACM.
- [23] Kock, N. (2011). Media naturalness theory: Human evolution and behaviour towards electronic communication technologies. In *Applied Evolutionary Psychology* (p. Applied Evolutionary Psychology, Chapter 0023). Oxford University Press.
- [24] Lent, B. (2012). Cybernetic Approach to Project Management: Where Sense Making Intelligence Is Needed. In *Autonomous Systems: Developments and Trends* (pp. 97-108). Springer Berlin Heidelberg.
- [25] Manstead, A., Lea, M., & Goh, J. (2011). Facing the future: Emotion communication and the presence of others in the age of video-mediated communication. In A. Kappas & N. Krämer (Eds.), *Face-to-Face Communication over the Internet: Emotions in a Web of Culture, Language, and Technology* (Studies in Emotion and Social Interaction, pp. 144-175). Cambridge: Cambridge University Press.
- [26] Matarazzo, G., & Sellen, A. (2000). The value of video in work at a distance: Addition or distraction? *Behaviour & Information Technology*, 19(5), 339-348.
- [27] McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, 60, 48-58.
- [28] Miller, D., & Sinanan, J. (2014). *Webcam*. Cambridge; Malden, MA: Polity Press.
- [29] Misra, S., Cheng, L., Genevie, J., & Yuan, M. (2016). The iPhone Effect. *Environment and Behavior*, 48(2), 275-298.
- [30] Muhlfelder, M., Klein, U., Simon, S., & Luczak, H. (1999). Teams without trust? Investigations in the influence of video-mediated communication on the origin of trust among cooperating persons. *Behaviour & Information Technology*, 18(5), 349-360.

- [31]Neustaedter, C., & Greenberg, S. (2012, May). Intimacy in long-distance relationships over video chat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 753-762). ACM.
- [32]Parkinson, B., & Lea, M. (2011). Video-linking emotions. In A. Kappas & N. Krämer (Eds.), *Face-to-Face Communication over the Internet: Emotions in a Web of Culture, Language, and Technology* (Studies in Emotion and Social Interaction, pp. 100-126). Cambridge: Cambridge University Press.
- [33]Przybylski, A., & Weinstein, N. (2013). *Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality*. *Journal of Social and Personal Relationships*, 30(3), 237-246.
- [34]Ruff, R. M., Niemann, H., Allen, C. C., Farrow, C. E., & Wylie, T. (1992). The Ruff 2 and 7 selective attention test: a neuropsychological application. *Perceptual and Motor Skills*, 75(3_suppl), 1311-1319.
- [35]Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, 43, 450-461.
- [36]Ryan, R., Rigby, M., & Przybylski, C. (2006). The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion*, 30(4), 344-360.
- [37]Self-assessment Grids (CEFR). (n.d.). Retrieved May, 2018, from <https://www.coe.int/en/web/portfolio/self-assessment-grid>
- [38]Shin, S. Y., Jang, J. W., & Bente, G. (2017). The benefits of distance and mediation: How people react to conflicts in video chat vs. FtF. *Computers in Human Behavior*, 73, 1-8.
- [39]Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. London [etc.]: Wiley.
- [40]Simons, D., & Chabris, C. (1999). Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events. *Perception*, 28(9), 1059-1074.
- [41]Storck, J., & Sproull, L. (1995). Through a Glass Darkly What Do People Learn in Videoconferences? *Human Communication Research*, 22(2), 197-219.
- [42]Subrahmanyam, V., & Ravichandran, K. (2013). Technology Online Distance Mode of Learning. *International Journal of Humanities and Social Science Invention*, 2(1), 5-13.
- [43]Tania Rascia. (2017, December 18). How to Use Local Storage with JavaScript. Retrieved from <https://www.taniarascia.com/how-to-use-local-storage-with-javascript/>
- [44]Tremblay, M., Blanchard, C., Taylor, S., Pelletier, L., Villeneuve, M., Irving, Greg, . . . Latham, Gary P. (2009). Work Extrinsic and Intrinsic Motivation Scale: Its Value for Organizational Psychology Research. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, 41(4), 213-226.

- [45]Tu, C. H., & Mclsaac, M. (2002). The relationship of social presence and interaction in online classes. *The American journal of distance education*, 16(3), 131-150.
- [46]Turkle, S. (2016). *Reclaiming conversation: The power of talk in a digital age*. NY, NY: Penguin Books.
- [47]Turkle, S. (2017). *Alone together: Why we expect more from technology and less from each other*. Hachette UK.
- [48]Walsh, D. (2016, September 23). Camera and Video Control with HTML5. Retrieved from <https://davidwalsh.name/browser-camera>
- [49]Walther, J. (2011). Visual cues in computer-mediated communication: Sometimes less is more. In A. Kappas & N. Krämer (Eds.), *Face-to-Face Communication over the Internet: Emotions in a Web of Culture, Language, and Technology* (Studies in Emotion and Social Interaction, pp. 17-38). Cambridge: Cambridge University Press.
- [50]Walther, J. B. and Bazarova, N. (2008). Validation and application of electronic propinquity theory to computer-mediated communication in groups. *Communication Research*, 35, 622–645.
- [51]Wiener, M., & Mehrabian, A. (1968). *Language within language : Immediacy, a channel in verbal communication* (The Century psychology series). New York: Appleton-Century-Crofts.
- [52]Wilson, H. K. (2015). A Critical Evaluation of Selective Attention Measures. *Honors College Capstone Experience/ Thesis Projects*. Paper 538.
- [53]Wilson, T.D., Reinhard, D.A., Westgate, E.C., Gilbert, D.T., Ellerbeck, N., Hahn, C., Brown, C.L. and Shaked, A. (2014). Social psychology. Just think: The challenges of the disengaged mind. *Science (New York, N.Y.)*, 345(6192), 75-7.

8. APPENDICES

1. Experiment

8.1.1. The edited script of the simulated VMC environment

Original script texts from Simon Corcoran are underlined, and the script file given to the speaker was without any numbered item and marks. (The speaker is free to add notes herself after getting the script.)

Instructions to the speaker:

- Pretend like you are actually talking to someone through the webcam
- Speak at your normal speed

-----start-----

Hi there, can you hear me? (pause) Okay, great! It's nice to meet you. Well, (1) I'm going to talk about an interesting conversation that (2) I had a couple of weeks ago in a music shop. (3) I was walking along one of the main shopping streets in the city centre, when a large window displaying all sorts of musical instruments caught my eye. Out of curiosity, (4) I decided to go in and have a look around.

The person (5) I ended up speaking to was a shop assistant on the second floor, in the area of the shop dedicated to acoustic guitars. (6) I hadn't intended to speak to anyone, but the assistant approached me in a friendly way and asked whether (7) I had any questions.

(8) I explained to the assistant that (9) I hadn't played the guitar for years, but that (10) I wondered what the differences were between the various acoustic guitars on show. He talked to me about the different makes and models, whether they were factory or hand made, the woods and varnishes used, the variation in sound quality, and of course the price range.

(11) / found the conversation fascinating because the shop assistant was so knowledgeable. It was obvious that he had a passion for the guitar, and he didn't mind talking to me even though (12) / had made it clear that (13) / didn't intend to buy anything. He even picked up and played three or four of the instruments to demonstrate the differences in their sound.

So that was the whole story, wasn't it great? Anyway, it was nice talking to you, have a good day! (14) /'m gonna hang up. Bye!

-----end-----

Timeline of the video call is as below (format: *(no.) seconds,milliseconds*), the time between 12th and 13th is the minimum one among all stimulus time gaps.

- (1) 14,240
- (2) 18,050
- (3) 21,270
- (4) 33,100
- (5) 37,290
- (6) 46,180
- (7) 53,030
- (8) 55,170
- (9) 57,170
- (10) 60,240
- (11) 80,240
- (12) 93,040
- (13) 94,240
- (14) 113,280

2. Survey

The online survey was build with Google Form (<https://goo.gl/forms/t5GI6XN5B3PsMBL62>); survey sample below is a pre-filled version.