Getting into the Meeting-Feeling: An Explorative Analysis of Presence in Videoconferencing

Shirley A. Elprama[1], Katriina Kilpi[1], Aljosha Demeulemeester[2], An Jacobs[1], Peter Lambert[2] and Rik Van de Walle[2]

[1]iMinds-SMIT-VUB, Vrije Universiteit Brussel, Pleinlaan 9, B-1050 Brussels, Belgium
[2]iMinds-MMlab-Ugent, Ghent University, Department of Electronics and Information Systems, Gaston Crommenlaan 8 bus 201, B-9050 Ledeberg-Ghent, Belgium
{shirley.elprama@vub.ac.be, katriina.kilpi@vub.ac.be, aljosha.demeulemeester@ugent.be, an.jacobs@vub.ac.be, peter.lambert@ugent.be, rik.vandewalle@ugent.be }

Abstract

People working at different locations still choose to meet face-to-face despite the availability of videoconferencing technology. A plausible explanation is the lack of social and physical presence – the feeling of being together inside a meeting space. After reviewing literature on presence, a multidisciplinary team first developed a mockup, followed by a prototype of an innovative videoconferencing system consisting of: (1) a Virtual Meeting Room displaying all participants as avatars and (2) video streams of all participants. By conducting focus groups and expert reviews, we evaluated which elements affected the sense of presence. Also, we argue that presence in formal videoconferencing meetings is not always necessary depending on various factors, such as the relevancy of the meeting topic to a participant.

1. Introduction

Regardless of the array of videoconferencing tools available, people still choose to come together in face-to-face meetings. Next to technical problems occurring in videoconferences and the prestige associated with travelling for work (Roy & Filiatrault, 1998), several presence related reasons are given. These include: lack of body language (Teoh, Regenbrecht, & O'Hare, 2010), lack of eye contact (Bekkering & Shim, 2006), and the overall lack of social contact (Panteli & Dawson, 2001).

Feeling presence in a videoconferencing meeting is important for a number of reasons. For instance, heightened sense of presence can translate into increased enjoyment, involvement and task performance efficiency (Lombard & Ditton, 1997; Witmer & Singer, 1998). Since immersive virtual environments can be distinguished from regular desktop systems because they "*afford a sense of presence*" (Slater & Wilbur, 1997, no page), Slater and Wilbur (1997) suggest that finding the important factors that contribute to presence can guide the future of technology. In this paper, we report how our multidisciplinary team explored which factors could be translated into a new technological concept: the ICOCOON system.

The ICOCOON project aims at creating an immersive teleconferencing experience without the need for dedicated and expensive videoconferencing hardware, making use of multiple cameras, sensors, video directing and mixed virtual/real representations of the participants. The vision was to create an immersive teleconferencing system (iMinds, n.d.), making use of a virtual environment where all participants are depicted as photorealistic avatars. In this technological project there was room for multidisciplinary work, integrating a conversation between the research of a small team of social scientists (three of the authors) and different teams of different specialties on the technical side. From the abundance of possible technical solutions, however, collaborative design decisions had to be made in order to explore whether these decisions worked in favor of immersion or not.

Although the concept of immersion has often been used interchangeably with presence (McMahan, 2003), we argued within the project that immersion is just one aspect of presence and that the focus should be broader. In this we follow Witmer and Singer (1998) who describe presence: *"... as the subjective experience of being in one place or environment, even when one is physically situated in another*" (p. 225). They describe immersion as "*a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences*" (Witmer & Singer, 1998, p. 227).

Therefore, the research question was reframed into the following: of all the technical virtual environment options that are possible in such a new teleconferencing system, which work in favor of presence and which against it?

To be able to research this question as well as to provide input during the implementation phase, we developed a video mock-up to be used as a probe. This mock-up consisted of a Virtual Meeting Room (VMR) showing all meeting participants as avatars, while also displaying video streams of all of the meeting participants. Although we are acquainted with certain presence enhancing aspects from previous studies in existing literature, designing a system, which would integrate these aspects does not necessarily imply that the actual system as a whole will create a feeling of presence.

By giving the potential end-users a holistic experience of the concept via the focus group interviews where they got to experience the mock-up system, we were able to give advice on the presence enhancing factors as well as on other aspects, which are beyond the scope of this paper.

This advice was implemented and the working prototype was developed. This prototype was evaluated by four user research experts, two of which who were not involved in the previous phases of the project.

Before we report on our own findings, we begin section 2 describing our starting point, the existing research findings on presence and its determinants by other scholars. Section 3 provides an overview of the ICOCOON videoconferencing system. Section 4 describes the methods, namely, how focus groups and expert reviews were used to gather data about the way presence is perceived or understood. Section 5 outlines the findings of the focus groups and the expert reviews. Lastly, this paper concludes with a discussion of the limitations of the study and the concluding remarks.

2. Dimensions of presence

To determine the elements of presence that can be experimented with when designing a new system, we examined existing literature on presence. Obviously, the basic definition, as stated in the introduction, should be more refined to understand the complexity of the concept. This section describes concepts fundamental to this piece of research.

Central in understanding the phenomenon of presence is both its social ("*the feeling of being together (and communicating) with someone*") and physical component ("*the sense of being physically located somewhere*") (Freeman, 1999; cited by IJsselsteijn, Ridder, Freeman, & Avons, 2000). We already mentioned the precondition of a state of immersion, but another essential component to experiencing presence is involvement. It is described as "*a psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli or meaningfully related activities and events*" (Witmer & Singer, 1998, p. 227). To maintain involvement in everyday life, however, is different because our focus is continually shifting while performing various tasks.

Similar to Bleumers, Van Lier and Jacobs (2009), who used Lombard and Ditton's (1997) dimensions of presence to identify which aspects of presence are promoted in three virtual worlds, we use these dimensions to investigate the way in which they can be incorporated into the design of the ICOCOON system. Subsequently, we briefly explain four relevant dimensions and illustrate the way theory was applied to the design of the mock-up and the prototype of ICOCOON. However, first we will present a brief description of the ICOCOON system to understand the concept and illustrate what the users see.

The ICOCOON system consists of two components: 1) a video stream showing users and 2) the VMR that represents these users from various locations as avatars in a virtual environment sitting around a virtual table (see Figure 1). When users join a meeting, they are visualized as an avatar sitting at a single table in the virtual environment. This table combines all of the participants at all of the locations around one table. In order to indicate whether participants are from the same physical meeting location, the parts sections of the table that are in close vicinity to these participants would turn the same color (see figure 3). Based on predefined actions, the avatars are able to mimic actions such as sitting, walking, presenting and raising their hands.

2.1. Presence as social richness

Lombard and Ditton (1997) introduce social richness as the first conceptualization of presence. However, like Bleumers et al. (2009), we have instead chosen to adopt the definition of social presence as described by Biocca, Harms and Burgoon (2003). Biocca et al.'s conceptualization distinguishes three (interrelated) dimensions of social presence: 1) co-presence (i.e. users becoming aware of another actor within range); 2) psychological involvement (i.e. establishing a relationship with the other actor) and 3) behavioral engagement (i.e. picking up (non-) verbal cues from the other actor that enable mutual understanding).

Co-presence can be encouraged by visualizing or having a visual representation of others, for instance on video or as avatars. Moreover, seeing avatars in a virtual environment facilitates presence (Bailenson, Yee, Merget, & Schroeder, 2006). Furthermore, Ratan, Santa Cruz and Vorderer (2007) found empirical evidence suggesting that similarity between avatars and the persons they are representing significantly affected presence. Finally, "*the knowledge that one is also being viewed should further establish a sense of co-presence*" (Beers Fägersten, 2010, p. 183).

For the design of ICOCOON, it was chosen to portray the users as avatars in a VMR. This way, the user is aware of the presence status of all users in the virtual room. In fact, similar to face-to-face meetings, the users and experts could constantly be aware of other participants' presence status, i.e. who was in the room and who had just stepped out. In the mock-up and in the prototype, one's presence status was visualized by showing meeting participants on the video stream and also by showing them as avatars in the VMR. Furthermore, by allowing for verbal interaction between users, they can engage in a conversation with other meeting participants and build a relationship with them (this was only implemented in the prototype, not yet in the mock-up). Since some (non-) verbal cues might get lost in translation from users to avatars, behavioral engagement in this setting is ensured by having such cues as raising a hand to indicate an intention to speak, visualized.

2.2. Presence as realism

A second conceptualization of presence is defined as a sense of realism, as delineated by Lombard and Ditton (1997). Within realism, the authors identify two types of realism: perceptual realism and social realism. In addition, Bleumers et al. (2009) added '*realism of interaction*' as a third type of realism.

Perceptual realism refers to the degree to which a medium produces accurate representations, as if that is, the degree to which they "*look, sound, and/or feel like the "real" thing*" (Lombard & Ditton, 1997, p. 4).

At this stage of the research, it was not yet possible to portray participants as realistic avatars, as is the end goal of the ICOCOON project. Therefore, elements of perceptual realism were not incorporated in either the mock-up or in the prototype. But the video stream did portray users in a realistic manner. Furthermore, in the prototype, a simple phone connection was used in order to mimic sound, which enabled a high quality audio connection.

Social realism (or plausibility) refers to what does or could happen in real life. In other words, an event is considered plausible, when it fits a person's prior knowledge and experience (Connell & Keane, 2006). Because the ICOCOON system would be used for videoconferencing meetings, we simulated such a meeting with the mockup and the prototype.

Realism of interaction refers to the sense that the interaction between objects and people represented by the medium, resembles real world interaction (Bleumers et al., 2009). For instance, ICOCOON allows users to interact with each other verbally through discussion and dialogue as is common in real life.

Furthermore, technical difficulties such as delays should be avoided, since they are harmful to the realism of the interaction and, in turn, can cause breaks in presence (Beers Fägersten, 2010).

2.3. Presence as transportation

Transportation is the third dimension defined by Lombard and Ditton (1997). They use this term to illustrate the way in which presence can transport users to a shared space. This can happen by transporting one user to another place; transporting a place and its objects to the user; or by transporting multiple users to a shared place (Lombard & Ditton, 1997).

By integrating a VMR with a round table in the design of ICOCOON, our aim was to transport multiple users to a shared environment. Moreover, it has been suggested that placing multiple users in a shared virtual environment creates a feeling of presence (Bregman & Haythornthwaite, 2003; Heeter, 1992).

Another way to encourage transportation is to use a first person perspective and through this, the perception of self-movement (Witmer & Singer, 1998). A way to do this is by visualizing participants as avatars engaged in a group discussion, as was done by Hoffmann, Klatt, Lam-chi, Haferkamp and Krämer (2010) who investigated decision-making in virtual environments. Avatars in their study were displayed from a first person or a third person perspective. Their results suggest that participants who were visualized from the first person perspective, reported feeling more involved in the group discussion. In turn, it has been suggested that more involvement plays a central role in the development of feeling of presence (Witmer & Singer, 1998).

Finally, accepting that one is in a virtual world is partly determined by having control over it (Heeter, 1992). When users have more control over the virtual environment, they also experience more presence

(Sheridan, 1992; cited by Witmer & Singer, 1998). One way to exercise control over the virtual environment is to be able to interact with it. Slater and Wilbur (1997) describe how interactivity, being a part of shared immersive technology, refers to the extent to which users can create and modify the content in the shared environment. Interactivity, however, is a significant part of videoconferencing and also an important part of creating a feeling of presence, given the fact that a lack of it can reduce presence (Tu & McIsaac, 2002).

2.4. Presence as immersion

The fourth dimension of presence is immersion (Lombard & Ditton, 1997). This dimensions consists of two components that Bleumers et al. (2009) name sensory and attentional immersion. While the sensory component "*refers to a state in which the user's senses are fully addressed by the medium*" (Bleumers et al., 2009, p. 2), the attentional component "*refers to a state in which the user's attention is entirely devoted to the mediated content.*" (Bleumers et al., 2009, p. 2). The two components are intertwined since visuals and audio can engage users' senses, which in turn can catch a user's attention.

The attentional component of presence is sometimes also referred to as involvement, a concept described earlier in this paper. Witmer and Singer (1998) state that it is difficult to maintain involvement in everyday life, because our focus is continually shifting while performing tasks.

Witmer and Singer (1998) found that the level of involvement depends on how significant the stimuli, activities or events, are deemed by the individual. They reason that increased focus on stimuli in the virtual environment leads to more involvement in the experience, which, in turn, leads to an increase in presence in the virtual environment.

When attention is diverted, presence can be threatened. Waterworth & Waterworth (2001) distinguish two ways attention can be diverted. The first one concerns the situation when the attention is shifted from the virtual environment to the external environment. In our case, this would mean a shift in focus from the system (video stream and VMR) to the meeting participants present in the same (physical) room.

The second way is in a situation when a person's attention is neither focused on the virtual or on the external environment. An example of this is when a meeting participant's mind would wander off during a meeting, while not paying attention to the video stream or the VMR. Another reason not to pay attention can be because a meeting participant is multitasking, which often happens in meetings (Mark, Grudin, & Poltrock, 1999).

2.5. Presence within medium

Although Lombard and Ditton (1997) defined a fifth and sixth conceptualization of presence, i.e. presence as social actor within medium and presence as medium as social actor, we will not discuss it here. Social actor within medium is not relevant, since we did not foresee users interacting in a social manner with parts of the system, even in situations when interaction was not possible. Furthermore, in our work we concentrated on presence within ICOCOON, and not on the medium itself as a social actor. Therefore, presence as social actor and as medium are not applicable in our case.

3. The ICOCOON videoconferencing system

ICOCOON is a videoconferencing system for professional use which aims at creating an immersive meeting experience, while at the same time being easy to set up, affordable and allowing flexible communication. The system consists of a video stream showing a limited amount of users that are selected by the system based on visual cues (see 3.1.) and the VMR, which presents the collected meta-data of all users from various locations as avatars in a virtual environment sitting around a virtual table (see 3.2.). The technical explanation of the VMR is explained in more detail in Demeulemeester et al. (2012).

The system consists of multiple smart camera networks, intelligent understanding of the scene, 3D capturing, 3D face animation, efficient information encoding and transport, and innovative rendering and displaying techniques.

The output of the system is a stream of meta-data that contains the locations of individuals in each room, a rough estimate of their shape and pose, and the direction in which they are facing. To lower the bandwidth requirements, smart cameras are used as they process information locally rather than sending video streams, and enable low latency solution. The gathered meta-data supports the creation of the two views that are presented to the user.

3.1. Video streams of all locations

Each installed smart meeting room comprises a calibrated smart camera network and one video aggregator, both responsible for extracting high and low-

level visual cues out of the captured video streams. Examples of extracted visual cues per meeting room location include the location and identification of people, gaze direction, hand gestures and activity status (sitting, walking, presenting, hands up/down). The virtual director is the central component of the system that analyzes all received visual cues from all locations and composites the output video stream, based on an inherent decision tree. This decision tree comprises a rule engine indicating how visual cues overrule each other.

3.2. VMR

The collected meta-data is also used to control the animations of 3D avatars of all users in a virtual environment (VE). In this context, the meta-data is seen as knowledge on the state of the physical world (i.e., the joining meeting rooms). The VE is visualized as a meeting room where all avatars are located. When participants join a meeting, they are visualized as an avatar sitting at a single table in the virtual environment. This table combines all the participants from all locations around one table. To indicate whether participants are from the same physical meeting location, the sections of the table that are in close proximity to these participants would turn the same color (see figure 3).

All relevant events in the meeting context have an associated animation or visualization in the VE. The VMR subsystem itself consists of the visualization server and a visualization client per user. The server adapts the virtual world state to reflect the state of the physical world and distributes this virtual state to the clients. The clients are running on the personal computing device of the users and each render their own unique view on the virtual world.

The VMR application was implemented using a 3D game engine. These are highly optimized to run a wide range of commodity hardware and provide all the required features for the VMR: a rendering engine for 3D graphics, scripting of object behavior and appearance, data driven avatar animation and networking support for state distribution. The game engine also allows deployment to a range of devices (e.g., PC, Mac, Android, etc.).

4. Methodology

In order to enable efficient collaboration among the multidisciplinary group of researchers, we used a scenario as a method to communicate among technical and social researchers. This scenario was created in collaboration with technical and social researchers. It served as the basis for the script of the video that was filmed, which in turn, resulted in a mock-up of the system. In other words, we created a prototype simulating the functionality of the videoconferencing system by means of a video.

4.1. Focus groups with video mockup

In order to be able to gain insights into a future technology in development and due to the fact that a working prototype was not yet available, we used a mock-up of the technology being developed. This allowed us to discuss the system with users in focus groups in the early stage, making it easier to implement their comments and suggestions in the working prototype.

The goal of the mock-up was threefold: 1) to gain insights into visualizations that are preferred by users (reported in Demeulemeester et al., 2012); 2) to probe topics such as privacy and trust in technology (reported in Kilpi, Elprama, & Jacobs, in press); 3) to gain an overall understanding of what the users thought of the system; in particular, we focused on which of the system elements might contribute to or threaten presence. This paper is limited to the findings that deal with presence.

A total of 23 people participated in six focus groups (mean group size = 3.8). Their ages ranged between 22 and 50 years old (average = 33 years). The users came from different organizations and had different professional backgrounds (such as a researcher, a manager and a historian). With the exception of two users, all others had experience with both teleconferences and videoconferencing. Users with experience in videoconferencing were purposefully chosen so as to be able to use their experiences with other video conferencing systems as the starting point for evaluating the strengths, weaknesses and possibilities of the proposed system. Each focus group lasted on average for 1,5 hrs (range: 1 hr 20 min - 2 hrs).

During each focus group, all users and two researchers sat around a table. This setup imitated one of the participating locations of a videoconference. The other parties in this simulated videoconference were shown participating in the meeting in a movie clip of circa 5 minutes consisting of three elements (see Figure 1): 1) remote participants of the simulated meeting; 2) the VMR with all meeting participants portrayed as avatars; and 3) presentation slides used in the simulated videoconference.

In the movie clip, a marketing agency's internal meeting was depicted, with participants joining in from different countries. While the video stream of the meeting transmitted the actions of the meeting participants, the VMR showed how these actions were translated to the avatars moving in the meeting room. For instance, events

such as entering the room and speaking were portrayed by avatars.

Each focus group followed the same procedure. First, the users received an explanation about the system. Because it was technically not possible to allow more interaction, the users were asked to pretend to be meeting participants in this simulated videoconference. Second, the users watched the entire movie clip simulating a meeting. While showing the movie clip, one researcher spoke and interacted with the participants (in the movie clip) of the simulated meeting. The other researcher observed and took notes. Third, the users were shown the movie clip scene by scene so as to discuss two to four design variations for events (e.g. walking into the room). These scenes were used to stimulate discussion among the users. Also, after seeing the clip, they were asked to rate how much they liked certain design variations (Demeulemeester et al., 2012). Finally, a topic list was used to guide the discussion. The researchers asked questions such as "*How important it is for you to feel that you are in a meeting with the remote participants?*" and " *Does it give you the feeling that you are in the same meeting?*"

Although the mock-up was by definition limited, it was useful as a probe to start the discussion about the elements of the system that threaten or contribute to presence.

The focus groups were recorded and transcribed and the transcriptions were subsequently coded using open coding.

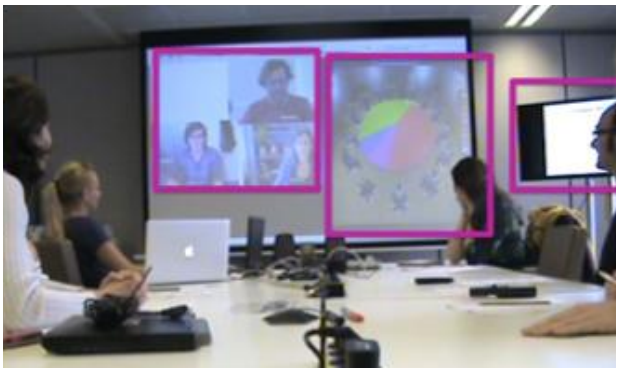

Figure 1. Focus group with three screens (from left to right): video stream, VMR and presentation slides.

4.2. Expert reviews with working prototype

When a working prototype of the system became available, we conducted two-expert (n = 3 and n = 4) reviews by having a videoconference meeting between two locations (Figure 2). Due to time limitations, it was not possible to conduct a review of the working system with users. In each location, all experts sat around a table.

Each expert sat in front of a laptop showing the video streams, the VMR and a combined feedback image of all participants of their location. The video streams of the working prototype consisted of separate video streams of each expert. The video stream of each participant was then automatically displayed in front of a virtual background. An ordinary phone connection was used to simulate an audio connection between the two locations in view of the fact that audio was outside the scope of this project.

The system was discussed during the review, which was based on a prepared topic list that listed subjects that included but were not limited to presence. Two experts (both specialized in virtual world applications) who were not involved in the project were interviewed individually after the review, in order to gain insights into their individual views on the system.

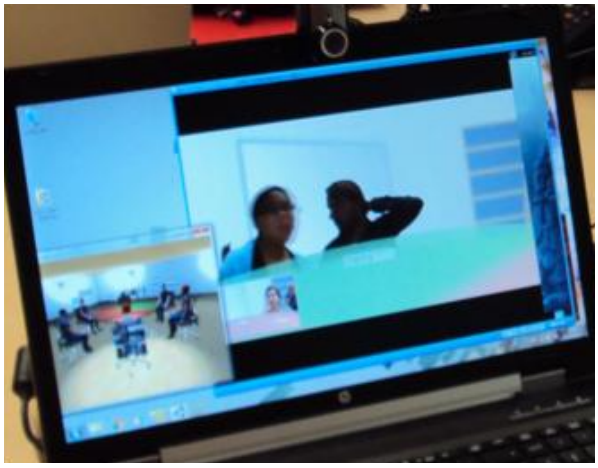

Figure 2. During the expert review, each expert had a laptop showing the VMR, the video stream from the remote location and a small feedback video stream of their own location.

5. Results

The participants did not feel immersed while viewing the technology mock-up. All experts who reviewed the working prototype, however, reported being able to achieve a feeling of presence during the expert review. Presumably, this is largely due to the possibility to interact with one another verbally and also viewing each other on the video

stream, which was possible in the working prototype. Nevertheless, both users and experts had an opinion on determinants of presence.

In order to illustrate the users' points of reference, we commenced the results section with a description of the way users perceived and defined presence. Then, while discussing the mock-up and the prototype, we discuss four conceptualizations of presence and the extent to which elements of ICOCOON contributed to or threatened the feeling of presence according to the users and experts. Note that some elements in the ICOCOON system could be translated as multiple dimensions of presence. However, we chose to address each element of the system together with the dimension of presence to which it was most closely linked. For instance, although the virtual table in the VMR was created to transport users into a shared environment, this element also touches upon realism, since it is common in meetings to sit around a table.

Names of all users are fictional and only represent the person's gender and age.

5.1. References to presence

Users described presence in various ways. Although they were not explicitly asked to define presence, the subject was brought up by researchers in discussions about the ICOCOON system and when sharing experiences with other videoconferencing systems.

Due to the level of abstraction of the concept, none but one user used the word "*presence*" explicitly. Some users that were familiar with the concept of immersion (through their work in virtual environments) used the words "*immersion*" or "*immersiveness*" during the focus groups. When describing the ICOCOON project goals, the researchers mentioned presence and immersion. It is therefore possible, that the focus group participants overheard and picked up terms they otherwise would not have used.

In general, the users did not seem to distinguish between the concepts of presence or immersion, but used them interchangeably (cf. McMahan, 2003). Nevertheless, the users seemed to have a clear view of what was meant with presence and immersion as can be read in the following.

5.1.1. Social presence. A user defined presence as the "*feeling of being together*" (Timothy, 37) suggesting that presence is a subjective experience of two or more people being together. Another user described presence through the action of being engaged in a discussion: "*Okay, yes. It did give the feeling of 'okay, we are having a discussion together*'" (Paul, 32). All these descriptions combined reflect Freeman's definition of social presence as described in section 2.

5.1.2. Physical presence. Another user described presence as a feeling of being there: "I think it's either totally immersive, you get the feeling that you're there, like the Cisco thing, or, to me at least, you're [working] at home" (Thomas, 26). In this description, the user stressed the aspect of place in presence, which is similar to Freeman's definition of physical presence in section 2.

5.2. Presence as social richness

In the ICOCOON system attention was paid to social presence, i.e. providing clues of presence of others in the room. However, these results suggest that prototype was more successful in achieving social presence than the mock-up.

5.2.1. Interaction. One of the researchers had a speaking role and pretended to interact with the people on the video in the mock-up. However, the users did not interact with the people acting in the mock-up. The following citation illustrates how the mock-up lacked interactivity and how this diminished social presence. "*It becomes like a television show because you can't really interact with it. And it's the same thing here. You have a high expectation and then 'Oh, I can't do anything, it's static'*" (Tom, 47).

In contrast, while testing the prototype, the experts reported to feel presence mainly due to the interactivity and the video streams of the other experts.

While testing the mock-up and the prototype, users and experts mostly paid attention to the video stream and little to no attention to the VMR. However, many participants appreciated the idea of the VMR as an awareness support tool by providing cues (e.g. behavioral engagement). Our users reported that it often happens in teleconferencing or larger videoconferencing meetings that one loses sight of who is speaking at the moment and who is present in the room.

5.2.2. Non-verbal language. One user argued that dedicated systems are the ideal type of teleconferencing since they facilitate eye contact: "*When you're sitting at the table and there's the screen and it really looks like there's four people sitting in front of you (...) and you can look them straight in the eye (...) it really feels like you're talking to people*" (Thomas, 26). However, during the focus groups, eye contact was not possible because the system was shown as a movie clip. Similarly, in the expert

reviews, exact eye contact was not possible since web cams were placed on top of each laptop. Therefore, this type of non-verbal language did not contribute to the behavioral engagement between participants. A few users felt that the lack of eye contact negatively affected their experience and indeed kept them from feeling presence.

5.3. Presence as realism

In various ways, by using avatars and showing the video stream, we attempted to achieve a sense of realism in the mock-up and in the prototype. While the video stream was more successful in this respect, other elements of the system such as inconsistencies between the virtual and the real world and technical problems made it more difficult to create a sense of presence.

5.3.1. Video stream. In the focus groups and during the expert reviews, it was evident that most attention was being paid to the video stream. The prototype in particular with the video stream of the other group of meeting participants at a different location in addition to an audio connection between the two meeting locations and effortless interactivity, created a sense of presence. Also, the separate web cam streams of two to three participants were combined to a single video stream showing experts surrounding a virtual table, which resulted in a mixed reality.

5.3.2. Avatars. When discussing the avatars, the users mentioned some factors, which they felt, were important in helping them to identify with an avatar. One user felt her hair would be an important factor and was more precise about what she feels was required: "*Facial features are not that important. Maybe the hair color, that's the most important I think. If my hair is blond or brown, then I don't want to be black. I don't want to be with black short hair. Nobody will think 'oh yeah, that's [Kelly]'. But my clothes or, yeah, a female body is enough I think*" (Kelly, 23). A few users mentioned that the clothing of the avatar should match that of theirs. Generally, the users were modest in terms of the minimum requirements for their avatars. In the end, it's most important that the avatar does its job: "*I think it just needs to be realistic enough to recognize somebody*" (Paul, 32).

One idea that came up in one of the focus groups was to use avatars and the VMR as a replacement for meetings participants who did not have access to a laptop with a camera. In this case, these participants could be represented by avatars to indicate their presence. As such, the importance of an avatar's appearance became apparent.

One user (Ruth, 33) emphasized that the similarity between avatars and meeting participants is important since it simulates the feeling of talking with a person.

In sum, although Ratan et al. (2007) suggests that similarity with avatars is important in creating the feeling of presence, the avatars in the mockup and in the working prototype, according to the users, were not similar and realistic enough to create this feeling.

5.3.3. Technical problems. The bugs and delays in the mock-up negatively affected presence (cf. Held & Durlach, 1992; cited by Witmer & Singer, 1998). These technical problems harm presence in two ways: by threatening the realism of the interaction and by diverting attention. Subsequently, we describe how presence was threatened by these technical problems.

For instance, synchronizing the speaking role of the researcher with the events in the video stream of the mockup did not work out perfectly, which also reminded the users that what they were seeing was not a real videoconference happening in front of them.

Audio and video in videoconferencing can be disturbed by delays. One particular user explained why delays in audio annoy her and how a different standard applies to image quality: "*The audio quality - that should really be impeccable. Because it becomes incomprehensible when there's interference and when you can no longer recognize the voices and the like. An average image quality, that you can deal with, as long as there is no lag - that is really disturbing. The fact that it's not super sharp disturbs me less. As long as you can identify facial expressions*" (Ruth, 33). This illustrates how the VMR or the video stream in the mock-up were not enough to establish a sense of presence. Visual and audio feedback was required in order to build a sense of presence.

During the expert reviews, the image of one of the experts froze on the video stream. At the same time, the participant in room B asked the participants in room A to get started with the meeting without knowing that room A was trying to solve a technical problem. Such technical difficulties threatened presence (cf. Beers Fägersten, 2010) by threatening the realism of the interaction and diverting attention.

5.3.4. Inconsistency between reality and virtual environment. The users and experts indicated several inconsistencies between reality and the VMR, which, according to them, made it more difficult to feel immersed

and perceive a connection between the two. This section illustrates with three examples how presence as realism was threatened.

Firstly, some users indicated that they were (or to have been) confused by the differences between the actions in the video stream and by the way these actions were portrayed in the VMR. For instance, one participant explained how he disliked the design choice of an avatar walking to the board in order to present because "*he doesn't do so in real life either*" (Paul, 32). Interestingly, the same participant did not mind the avatar being depicted standing up to present, while in the video stream the person was sitting, concluding from this that it has to do with the plausibility of the action.

Secondly, the sitting order of the avatars in the VMR did not correspond to the meeting participants in the video stream. "*And also the fact that they're all sitting together but if you check the table where they are sitting, they're not sitting in the same order. So you might be thinking, "okay, this one is this one*". *No, it doesn't mean that*" (Jack, 34).

Finally, an inconsistency related to the previous one, was the disconnection between reality and the video stream in the working prototype that was used during the expert reviews. For instance, in real life, when person A would turn towards person B, one would expect the same persons on the video stream to face each other. However, this was not yet technically possible. In other words, the mixed reality video stream displayed person A not facing person B, although this was evident in real life.

5.4. Presence as transportation

Although it was clear to users and experts that the VMR, the virtual table within the VMR, and the first person perspective aimed at creating a sense of presence, opinions varied to what extent creating this feeling was successful.

5.4.1. Presence in VMR. During the focus groups, one of the users stressed how all of the interaction should take place in the VMR so as to create a feeling of presence: "*I think the interaction should then take place in the virtual environment, otherwise you'll never... A presentation should take place in the virtual [environment], otherwise it will never work*" (Ruth, 33).

The significance of Ruth's citation does not imply that all interaction should be taken in the VMR, but it does imply that a (design) choice should be made to have one screen as the focus, which can then facilitate the feeling of presence. The focus could be either the VMR or the video stream, or both (if they would be integrated).

This assumption was confirmed during the expert reviews, where most experts reported that they had paid most attention to the video stream. They perceived the video stream as the most important element that contributed to their feeling of presence. One expert did argue, however, that he felt that the VMR also added to the feeling of presence.

Surprisingly, we also noticed that a few experts focused on the video stream images, even during the time when a person located in the same room was speaking. These experts reported that especially in situations in which one might have to physically turn towards another participant in the same room to face them, they would rather choose to look at their image on the screen. If the participant was located in front of them, it was similarly convenient to look at them in real life than look at them on the video stream. Presumably, during the expert reviews, a feeling of presence was developed where all meeting participants were transported to the shared space and in which the participants of the shared physical space were neglected.

5.4.2. Virtual table. Several users and a few experts felt that the table in the VMR was beneficial for the feeling of presence. One of the users stated that what he observed "*created the feeling of being together in a meeting*" (Bob, 28). Another user (Kelly, 23) illustrated how the feeling of presence - or "*that round table feeling*" as she described it - could be shattered. She explained the way in which the table in the VMR stimulated the feeling of presence by connecting all remote participants and placing them as avatars around a virtual table. She believed, however, that this round table feeling would disappear if participants of each remote location would sit at different tables. Some users felt the table was merely a facilitator by enabling a situation in which presence is created. An example of this is the situation when participants (sitting around a table) are engaged in a discussion and through this action a feeling of presence is created.

However, some users disagreed on the table's role in creating a feeling of presence. For instance, one user argued the following: "*Because you have such a huge table, you don't have... If you sit around a table like that, you don't have a feeling of being together*" (Timothy, 37).

In line with this reasoning, the pie chart on the table was perceived as dividing participants from multiple locations instead of bringing them together: "*Don't put everyone from the same continent on the same side,*

because then you again split up people. (…) It's a meeting so you should have everyone in one room and now you create these... [divisions in the pie chart]" (Tom, 47).

Although most users identified the table as a factor in contributing to presence, it might be that the mere existence of the VMR in which avatars were present in a shared space facilitated presence (cf. Heeter, 1992).

5.4.3. First person view. During the focus groups, a video was shown of the VMR in different perspectives, such as a first person view and a top view (see Figure 3). According to a few users, this first person view facilitated the feeling of presence more than the top view.

"Julian (50): *I liked the change [in] the camera angle. I think it's nice. So [the first person view] is very nice for me.*

Timothy (37): *Except for the scene...*

Julian: *And very strange, it's a view you never get as a human.*

Timothy: *It was like from the floor, looking to the ceiling, somebody lying on the floor.*

Researcher: *But you still liked it?*

Julian: *Yes I liked it, it gives another view. I think there's the effect that you get out of your body and walk around."*

Seeing oneself move, might explain why some users preferred the first person view (cf. Witmer & Singer, 1998). However, not everyone agreed with this opinion. Others underlined that the first person view was "*confusing*" or "*distracting*". In fact, when the clip of the first person view was shown, there was a bug in the animation, which might explain why some users felt confused, even though the bug was explained to them.

5.4.4. Feedback. During the focus groups, some users emphasized that they missed a feedback image of themselves: "*And also I like to always see our picture, to see what other people might see*" (Jack, 34). They, however, did not indicate explicitly if this was important for the facilitation of presence, but rather that it made them feel more comfortable when seeing how one is perceived by others.

5.4.5. Control. The need to exercise control became apparent when some users hypothesized that they would feel more comfortable when having more control options in their hands, e.g. the possibility to zoom in on the video stream to see a specific person's reaction to something or having access to a mute button.

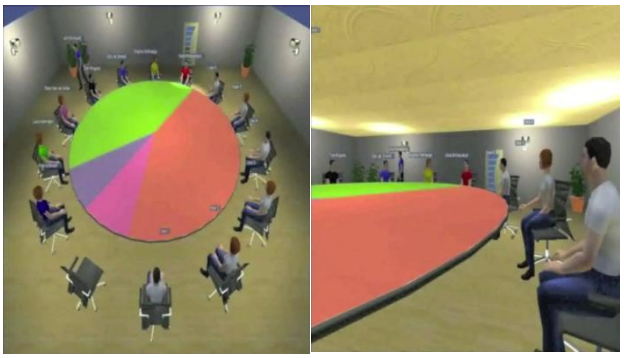

Figure 3. Top view (left) and first person view of VMR

5.5. Presence as immersion

Feeling immersed while watching the mock-up was not an easy task for the users in the focus groups. However, multiple users argued that feeling immersed is not always needed in videoconference meetings, since not all parts of meetings are relevant or simply because people like to multitask during videoconferences.

5.5.1. Multiple points of focus. A recurring complaint across all focus groups was that the mock-up consisting of three screens (presentation slide, video stream and VMR) provided too much information, which made it difficult for the users to know what to focus on. Although most users complained about the amount of information they were subjected to, they did not explicitly state this aspect as the reason for not feeling immersed. However, one user described the way the amount of information breaks the feeling of presence: "*Well, for me it's also like putting the right thing on the left side or the left side inside the right. I don't know which one I'd prefer; I'd have to see it but not two separate things. It's too complicated, it asks too much. (…) I have the feeling that it breaks the immersiveness*" (Peter, 32). In contrast, the experts mainly focused on the video stream and were less distracted by the other elements.

In addition, the setup requires that participants bring their laptops. One of the users expressed his concern regarding the way this would affect the feelings of presence: "*It's a bit strange to come to a meeting and everybody is on his PC*" (Rob, 33). Another user justified or explained explicitly why she was concerned: "*Then you kind of disconnect people who are sitting [together] because everybody's behind their screen*" (Anna, 34).

5.5.2. Low involvement due to relevancy. How easily a person disengages from a video mediated meeting has to do with how committed an individual is to the task

at hand, i.e. the level of involvement a person feels. This lack of involvement concerns the first type of diverting attention when attention is shifted from the virtual environment to the external environment (Waterworth & Waterworth, 2001). One user pointed out that videoconferences are often boring: "*You're in a certain a meeting and it's very boring because 70% of all conference calls are very boring. So you can actually walk away*" (Steven, 34). A few other users pointed out that not all meetings are always relevant from beginning to end: "*Not every part of the meeting is equally important for you. You might be invited to the whole meeting and only the last half an hour is important for you*" (Peter, 32). As stated by the user, not all parts of a meeting are relevant for all the members of the meeting. Thus, it is more likely that people will "*zone out*" (Thomas, 26), which means that this factor of presence cannot entirely be controlled by the design of the virtual environment.

5.5.3. Multitasking. On the one hand, several participants reported that multitasking happens when the meeting is not perceived to be relevant enough, when it has become too boring, or when other things are more important. They also argued that multitasking is a common practice in meetings (cf. Mark, Grudin, & Poltrock, 1999). "*I think it happens all the time, people without laptops scribbling, making noise, but it doesn't mean not paying attention*" (Steven, 34). On the other hand, some users were against multitasking. They found it mainly distracting; one participant even called it "*disrespectful*" (Heather, 26) towards others.

Because the ICOCOON system picks up only such predefined cues as someone speaking audibly or someone raising a hand, a participant has the possibility to multitask without disrupting the concentration of other people.

In sum, meeting relevancy and multitasking during meetings suggest that systems such as ICOCOON should design for immersion but also allow participants not to be immersed (cf. Lyons, Kim, & Nevo, 2010).

Even though at first users reported being distracted by all the information and by many focus points in the mock-up, by the time we tested the prototype with experts, the distracting elements had been toned down. The experts felt that the interaction was natural, and although the VMR did not appear to be useful in such a small meeting set up, it was not a distracting element either.

When users and experts reported feeling presence, it was predominantly because they kept following the video stream and hence felt presence via the video stream. This made it difficult to determine, which elements of the VMR could work in favor of presence. However, as presented in the results, we were able to get some indication of the elements of the VMR that should be further studied such as the table in the VMR, having more control over the environment and other elements. Integrating the information of the VMR and the video stream in a single prototype would also be relevant to explore, since this might facilitate a shared immersive environment without two or more competing sources of information.

Even though some users raised the point of the lack of eye contact and therefore – according to them - not being able to experience presence, we found that in the ICOCOON system, other things (such as interaction) took over and substituted for eye contact. In a similar vein, Isaacs and Tang (1993) found that participants using a desktop conferencing prototype called a person's name to indicate which speaker was next since the set up of the system did not allow one-to-one eye contact.

Similarly, users are sometimes more flexible about the amount of realism they expect to enable feelings of presence. For example, where one user reported disliking the avatar walking to the board to present because it was not realistic, the same person was not as strict about an avatar standing at his place to present while on the video stream the participant was actually sitting. This has most probably to do with the difference of action – presenting at ones place and walking to the board, which is another, explicit action. In contrast, although in reality people choose to sit down while presenting, the commonly agreed upon correct way to present is by standing up. Therefore, the VMR's depiction is plausible for the user and users are more flexible about the degree of realism they expect.

6. Discussion

Although the main goal of the ICOCOON project was to create an immersive video conferencing system, our findings suggest that feeling presence in a meeting context is not always required. For instance, depending on the meeting relevancy to the participant, people are not always paying attention during meetings and multitasking is common practice.

Furthermore, even though we tried to implement various factors that contribute to presence, we found that not all these factors were equally effective. This suggests that implementing multiple ways to encourage presence is important when designing an immersive system. Simultaneously, even though not all these factors were effective, the working prototype provided enough

interactivity and visual feedback to create a feeling of presence among the experts reviewing the system.

We acknowledge that the mock-up had limitations such as a lack of interactivity and visual feedback. Also, the lack of resemblance between the avatars and the users further hindered the possibility for an immersive experience. However, these limitations of the mock-up acted as probes, which elicited a lot of reactions and useful insights from the users.

Using a four interconnected conceptualizations of presence posed some challenges when trying to match these conceptualizations to elements of the ICOCOON system. Often, a certain element of the system could enhance presence based on more than one conceptualization. For instance, the lack of interactivity had not only an effect on social presence, but it presumably also was harmful for the transportation dimension. Future research could explore implementing various combinations of dimensions of presence to investigate which ones or which combinations of conceptualizations are more important. Also, when developing an immersive system, how many conceptualizations have to be in place in order for presence to be possible?

7. Conclusion

Previous research shows that the feeling of presence in videoconferencing meetings is important for enjoyment, task performance and efficiency. Using this as our guide, we explored ways to improve this feeling of presence in the ICOCOON project where we aimed at creating a new videoconferencing system making use of the new technological video and sensor opportunities.

In this paper we evaluated the translation of different presence related components with end users by developing a mock-up of the technology in progress, and the iterative result of our advice by an expert review.

Our first contribution is showing the translation of the different aspects of presence theory in a systematic way. It is very important to take into account the multiple facets at the same time, to enable a system that accounts for the individual variation in the experience of presence. Where for some presence means the physical actions replicated in the virtual environment (e.g. seeing themselves depicted by an avatar), others see it as the feeling of being engaged in a social action (e.g. discussion). For some, the combination of both is needed.

A second contribution is creating awareness of the fact that although presence can improve a meeting in general, it does not imply that the system should aim at creating a feeling of presence for every single participant. The goal of making an immersive teleconferencing system should not be treated as a holly grail to solve the current problems with presence in current non-high-end videoconferencing systems. Furthermore, rather than striving solely for technology mediated face-to-face meetings, it might be beneficial to estimate when these "accidental benefits" of videoconferencing do contribute to a successful meeting. For one, leaving the user in control of where he/she choses to multitask or zone out without any consequences from the system, is recommendable. The user should judge the meeting relevancy and his/her role in it.

Our final contribution is the approach to evaluate the translation of presence aspects in an iterative way. This iterative way of working allows us to have some impact on the technology under development and taking into account the system as a whole and not a singular aspect. As discussed in the previous section although there are limits to working with a mock-up, this method allowed us to simulate some of the main features of the technology being developed and create an experience to the users, we were able to provide a natural starting point for a discussion with the users. Furthermore, it allowed the user researchers to contribute to the development of the videoconferencing tool with user insights alongside the work of the developers. By working in an iterative manner, the prototype could be adapted incrementally, and the implemented recommendations to be investigated and evaluated.

Acknowledgements

We would like to thank Erwin Six and the Alcatel-Lucent Belgium team, Kris Naessens, Lizzy Bleumers and Koen Willaert for their contribution. Also, we thank the people taking part in the group interviews and the reviewers for giving feedback on an earlier version of this paper. The ICOCOON project is co-funded by iMinds, a research institute founded by the Flemish Government, with project support of IWT. Companies and organizations involved in the project can be viewed at (iMinds, n.d.).

References

Bailenson, J. N., Yee, N., Merget, D., & Schroeder, R. (2006). The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoperators and Virtual Environments*, *15*(4), 359–372.

Beers Fägersten, K. (2010). Using Discourse Analysis to Assess Social Co-Presence in the Video Conference Environment. In L. Shedletsky & J. Aitken (Eds.), *Cases on Online Discussion and Interaction: Experiences and Outcomes* (pp. 175–193).

Bekkering, E., & Shim, J. P. (2006). Trust in videoconferencing. *Communications of the ACM*, *49*(7), 103–107.

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a More Robust Theory and Measure of Social Presence: Abstract. *October*, *12*(5), 456–480. doi:10.1162/105474603322761270

Bleumers, L., Van Lier, T., & Jacobs, A. (2009). Presence and mediated interaction: A means to an end? In D. D. G. V. Donoso, D. Geerts, P. Cesa (Ed.), *Networked Television Adjunct proceedings of EuroITV 2009* (pp. 104–106). KULeuven.

Bregman, A., & Haythornthwaite, C. (2003). Radicals of presentation : visibility , relation , and co- presence in persistent conversation. *New Media & Society*, *5*(1), 117–140.

Connell, L., & Keane, M. T. (2006). A model of plausibility. *Cognitive Science*, *30*, 95–120.

Demeulemeester, A., Kilpi, K., Elprama, S. A., Lievens, S., Hollemeersch, C. F., Jacobs, A., Lambert, P., et al. (2012). The ICOCOON Virtual Meeting Room: a Virtual Environment as a Support Tool for Multipoint Teleconference Systems. *11th International Conference on Entertainment Computing* (p. 14).

Heeter, C. (1992). Being there: the subjective experience of presence. *Presence: Teleoperators and Virtual Environments*, *1*(2), 262–271.

Hoffmann, L., Klatt, J., Lam-chi, A., Haferkamp, N., & Krämer, N. C. (2010). Looking through the eyes of an avatar – How point of view influences the perception of virtual group discussions. *ICA 2010 Singapore*.

IJsselsteijn, W. A., Ridder, H. de, Freeman, J., & Avons, S. E. (2000). Presence: concept, determinants, and measurement. In B. E. Rogowitz & T. N. Pappas (Eds.), *Proceedings of SPIE* (Vol. 31, pp. 520–529). Spie.

Isaacs, E. A., & Tang, J. C. (1993). What video can and can't do for collaboration: a case study. *Proceedings of the Multimedia 1993* (pp. 199–205).

Kilpi, K., Elprama, S. A., & Jacobs, A. (in press). Exploring privacy and trust issues in a future immersive videoconferencing system. *Proceedings of the 46th Annual Hawaiian International Conference on System Sciences* (p. 10). Computer Society Press.

Lombard, M., & Ditton, T. (1997). At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, *3*(2).

Lyons, K., Kim, H., & Nevo, S. (2010). Paying Attention in Meetings: Multitasking in Virtual Worlds. *First Symposium on the Personal Web, Co-located with CASCON*.

Mark, G., Grudin, J., & Poltrock, S. (1999). Meeting at the Desktop: An empirical study of virtually collocated teams. In S. Bødker, M. Kyng, & K. Schmidt (Eds.), *Proceedings of the Sixth European Conference on Computer-Supported Cooperative Work* (pp. 159–178). Copenhagen, Denmark.

McMahan, A. (2003). Immersion, Engagement, and Presence. In M. J. P. Wolf & B. Perron (Eds.), *The video game theory reader* (Vol. Immersion,, pp. 67–86). Routledge, London & New York.

Panteli, N., & Dawson, P. (2001). Video conferencing meetings: Changing patterns of business communication. *New Technology Work and Employment*, *16*(2), 88–99.

Ratan, R., Santa Cruz, M., & Vorderer, P. (2007). Multitasking, Presence and Self-Presence on the Wii. *Proceedings of Presence 2007* (pp. 167 – 177).

Roy, J., & Filiatrault, P. (1998). The impact of new business practices and information technologies on business air travel demand. *Journal of Air Transport Management*, *4*(2), 77–86. doi:10.1016/S0969-6997(98)00009-X

Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence Teleoperators and Virtual Environments*, *6*(6), 603–616.

Teoh, C., Regenbrecht, H., & O'Hare, D. (2010). Investigating factors influencing trust in video-mediated communication. *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction - OZCHI '10* (p. 312). New York, New York, USA: ACM Press.

Tu, C., & McIsaac, M. (2002). The Relationship of Social Presence and Interaction in Online Classes. *The American journal of distance education*, *16*(3), 131 – 150.

Waterworth, E. L., & Waterworth, J. a. (2001). Focus, locus, and sensus: the three dimensions of virtual experience. *Cyberpsychology & behavior: the impact of the Internet, multimedia and virtual reality on behavior and society*, *4*(2), 203–13.

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators & Virtual Environments*, *7*, 225–240.

iMinds. (n.d.). iCocoon. Retrieved from http://www.iminds.be/en/research/overview-projects/p/detail/icocoon-2