

Carpe Diem: Exploring User Experience and Intimacy in Eye-Based Video Conferencing

Kaisa Väänänen-Vainio-Mattila¹, Katja Suhonen¹, Tina Gonsalves²,
Martin Schrader³, Toni Järvenpää³

¹Tampere University of Technology,
Unit of Human Centered Technology
P.O. Box 589
33101 Tampere, Finland
Telephone number, incl. country code
{kaisa.vaananen-vainio-mattila,
katja.suhonen}@tut.fi

²www.tinagonsalves.com
Po Box 395
Port Douglas, QLD, 4877,
Australia
0409 166 405
tina@tinagonsalves.com

³Nokia Research Center Tampere,
P.O. Box 00045
Finland
Telephone number, incl. country code
{martin.schrader,
toni.jarvenpaa}@nokia.com

ABSTRACT

Technology-mediated communication has become a prevalent means for remote communication between people. The mobile phone and video conferencing on PC enable broad bandwidth for everyday interactions. The effectiveness and social presence with traditional video conferences have been studied rather extensively. Our research arose from the partly artistic motivation of exploring how *intimacy* and emotional engagement between people can be supported by alternative types of image-based communication tools. We constructed a video conferencing system with near-eye displays, where – in addition to audio communication – the users could only see one of each others' eyes. The goal was to explore how such unusual eye-based display of the other person would make people feel during conversations. We conducted an explorative user study in a laboratory context with five pairs of users to understand their experiences with this system. The results show that this kind of mediated communication can cause a variety of experiences, such as “interesting”, “surprising”, “tranquil” or “pleasantly strange”. Overall, the full-screen display of the eye clears off distractions from the discussion and thus can help make the discussion more intimate and focused on the moment. We envision that a system like this could become a *tool for deeper listening*.

Categories and Subject Descriptors

H5.3. Information interfaces and presentation, Group and organizational interfaces.

General Terms

Design, Human Factors.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM Multimedia '11, Nov 28 – Dec 1, 2011, Scottsdale, Arizona, USA.
Copyright 2011 ACM 1-58113-000-0/00/0010...\$10.00.

Keywords

Mediated communication, video conferencing, user experience, art, eye-based communication, feelings, intimacy.

1. INTRODUCTION

Computer-mediated human communication has been an object of research in human-computer interaction (HCI) for decades. HCI studies have investigated PC-based and mobile communication systems using audio and video (e.g. [9]), and more recently other media such as haptics [5] [10]. In early phases the focus of video conference research was on task efficiency but with the emergence of user experience (UX) research, emotional qualities of systems have also been investigated (e.g. [1] [11]). At the same time, computer-mediated interactive art has made its way to art exhibitions but also to other types of events where people can try and experience various forms of situated installations, either alone or together with other people.

The motivation of this work arises from two directions. On one hand the research motivation comes from human-computer interaction (HCI) and more specifically, user experience (UX) research of new interactive technologies. The aim is to explore the hedonic qualities [7] of new technologies, in this case focusing on a system which may increase social connectedness which is one of the main hedonic values for users of new technology. On the other hand, the motivation of this work arises from the artistic motivation to study technology-based tools which disrupt habitual responses and cause emotional reactions which are “out of the ordinary”, creating new experiences.

McLuhan has stated “We shape our tools and then our tools shape us.” [14]. Today people are being connected in a rich variety of ways. A proliferation of new communication technologies such as live chat, mobile, video conferencing, text messaging and social networking have become embedded into our everyday communication, revolutionizing the way we share information and experiences to those close to us. In their book about loneliness, Cacioppo and Patrick [4] suggest that even though we are more connected than ever, people report feeling lonelier and more isolated than ever. Social communication interventions, including the one described in this paper, will foreground ideas of trust, intimacy and compassion while questioning the current directions

of mobile technologies: the trend of acceleration, 24/7 connectivity and daily data overload.

The artist's (one of the authors) work has always explored aspects of the intimacies and vulnerabilities of being human. As technology becomes ingrained in our every day, the aim of the interactive art is to explore what ways we can harness technologies as tools to enhance intimacy and disrupt habitual responses, lifting one out of their every-day, to elicit reflection, to create 'feeling'. The system built for this research is intended as a subtle intervention, disrupting usual conversation, seeing how this may have the potential to create more intimate, and more meaningful communication. It aims to create 'an awareness' in the participants, so they are more concentrated on the moment. The goal behind this explorative research was to find out more about the intuitive engagement of images to elicit felt experiences. The more long-reaching goal is to explore how affect, compassion and trust can be created using social communication tools.

We wanted to explore how, by reducing the visual input during remote communication, people would experience the discussions with people they know. Would they be able to better focus on the moment? Would they feel at ease with this kind of communication medium? As eyes of the communication partner are a central visual element in human-communication, we chose to base the system on seeing only one eye of the communication partner.

By building the prototype system for eye-based video conferencing we are not aiming at product development but to explore both the artistic and experiential qualities of alternative ways of human communication. The key design objectives were to aim at high immersion, out-of-ordinary and intimate communication experience. To this end, we built a prototype of such system and conducted a user study in the laboratory context with five pairs of users.

2.RELATED WORK

2.1.Video Conferencing User Experience

Previous user studies about video conferencing have focused on, for instance, the sense of togetherness and social presence in 2D and 3D video conferencing [9], and the comparison of high quality video conferencing, Skype calls with webcam and audio conferencing with respect to the experiential dimensions of emotional involvement, active participation, reciprocity, co-presence and group cohesion [1]. These studies have emphasis on the social aspects of video conferencing and conveying emotional information has been left to a minor role. IJsselstein et al. [11] have gone closer to emotions and examined the topic of connectedness and social presence in the context of media technology on theoretical level. According to them the media that enables transmission and display of nonverbal communicative cues, which are most visible in face-to-face communication, is considered as richer than the media where nonverbal channels cannot be used. They continue that nonverbal channels communicate information that is primarily affective in quality and connected with personal relationships, i.e. emotional information. Argyle and Dean [2] argue that intimacy in interpersonal communication is kept on an optimal level through factors such as

physical distance, smiling, eye contact and intimate topics of conversation. IJsselstein et al. [11] add to this list gestures, touching, vocal cues, turn-taking behavior in dialogues, the use of space, and verbal expressions directly acknowledging the communicative partner.

The use of gestures and touch to communicate intimacy or feelings through technology has been studied e.g. by Heikkinen et al. [10]. Research prototypes for this purpose include a ring that enables feeling your partner's pulse [17] and a haptic jacket for conveying hugs in teleconferencing [5]. The meaning of eye-contact has been studied in different communication contexts for decades as, for example, already in 1965 Argyle and Dean [2] examined the relation of eye-contact, physical proximity, and intimacy in communication. More recently Grayson and Monk [6] have studied mediated eye-contact in desktop video conferencing. They all agree that eye-contact is an important factor in interpersonal communication. Grayson and Monk [6] introduce some inventions that have been developed for preserving mutual eye-contact in videoconferencing, which is normally difficult to provide due to the disparity between the position of the camera and the position of the user's eyes on the screen. The solutions include for example using half-silvered mirrors [15] [12], or rotating video images of the participants so that they seem to look at each other [16].

There are some commercial systems which aim at preserving eye contact, such as See Eye2Eye¹, a teleprompter for webcams and Iris² videoconferencing system. However, there seems to be no related research on near-eye display –based systems focusing on users' experiences with those systems.

It appears that there are no earlier studies that would have involved mediated emotional communication with a partner in such a close visual proximity that you see only the other person's eye and nothing else. Close proximity in our study does not mean that the person would be physically very close, as for example in the studies by Argyle and Dean [2], but in the way that the user has a near-eye display only a couple of centimeters in front of his/her eyes and the view on the display is a close-up of one of the remote communicative partner's eyes. In this setup the eye-contact is almost forced – the users can still choose to close their eyes or look up, down, and to the sides, but the image of the partner's eye is dominating the view. The aim of this setup was to take a step further and force the participants to look at each other's eyes mediated through technology and study the whole experience with focus on intimacy and emotional communication.

2.2.Artistic Approaches

This work arises from the artist's interest in the positive social communication aspirations of *Fluxus*³ art movement (starting 1960). Fusing irreverence, playfulness, Da da and an interdisciplinary approach, Fluxus became an international network of artists with a primary aesthetic drive to integrate life into art, taking it out of the museums and galleries into everyday life. Yoko Ono, the best-known individual associated with Fluxus, tapped into a playful 'zen' approach, where the audience became

¹ <http://www.bodelin.com/se2e/>

² <http://www.redferret.net/?p=20807>

³ <http://en.wikipedia.org/wiki/Fluxus>

central to the process of art making. Ono's work often highlights the looking and thinking rather than the making.

Lygia Clark⁴ (1920–1988) a Brazilian artist, also championed the idea of the participation of the viewer. Clark investigated the language of the body and broke new ground in examining ideas of intimacy and new ways of communication. In 1967 she created *Sensorial Hoods*, an experiment involving eye pieces, ear covers, and a small bag that would be affixed over the participant's nose. The outcome of this experiment might be that a participant would use all their senses in a way they would not have thought possible.

The Distance Lab⁵ (in Scotland, no longer active) was a research organization that also looked into 'Slow' communication. They explored creative ways communication can be slowed down, and become a more personalized experience, supporting a sense of intimacy and closeness. Tomoko Hayashi's *Mutsugoto*⁶ is an intimate communication device intended for a bedroom environment. Instead of exchanging e-mail or SMS messages using generic interfaces in business-like venues, Mutsugoto allows distant partners to communicate in new ways. A custom computer vision and projection system allows users to draw on each other's bodies while lying in bed. Drawings are transmitted "live" between the two beds, enabling a different kind of synchronous communication that leverages the emotional quality of physical gesture.

In this work we aim to extend the look into deeper, more provocative ways of communication, relooking at how technology can elicit 'deeper' relationships with each other, and ourselves. From the user experience and HCI perspective we aimed at understanding how this new kind of communication system would be experienced by its users. To this end, we run an explorative study of five pairs using the system in a lab context.

3. THE SETUP OF EYE-BASED VIDEO CONFERENCING SYSTEM

The choice of the eye-based video conferencing system⁷ was driven by the goal of achieving high immersion and a strong focus on the possible role of the eye in creating intimacy and relaying emotions. To that end we chose a near-eye display that is covering the whole field-of-view of the wearer so that there won't be any other visually distracting input during the communication session. We also used an eye-looking camera to capture the eye of the wearer and transmitting it to the communication partner. We chose to greatly enlarge the eye of the communication partner in the video window (full screen mode) to maximize the possible emotional effect. The user could also see a small picture of their own eye in the bottom corner of the display. See Figure 1.



Figure 1. A pair sitting in separate spaces, using the eye-based video conferencing system. Each person sees a full-screen picture of the other person's right eye through the near-eye display. In addition, they can communicate through standard voice-over IP audio.

The wearable near-eye displays used in the experiments were prototype see-through displays with integrated eye gaze tracker functionality (developed by us, see [13]). The setup is illustrated in Figure 2. Compared to conventional near-eye displays with two sets of microdisplays and enlarging optics (one setup for each eye), the light from one microdisplay is split and passed to the eyes by using transparent plastic light guides. Angular field-of-view of the displayed image is 30 degrees diagonal, the virtual image focus is in infinity. See-through mode was not used as such as the transmission was blocked by adjustable liquid crystal shutters to increase the level of immersion. Still, the light guides enable flexible positioning of the eye-looking camera. The camera is located in front of the right eye, behind the light guides on the left side of the visual axis.

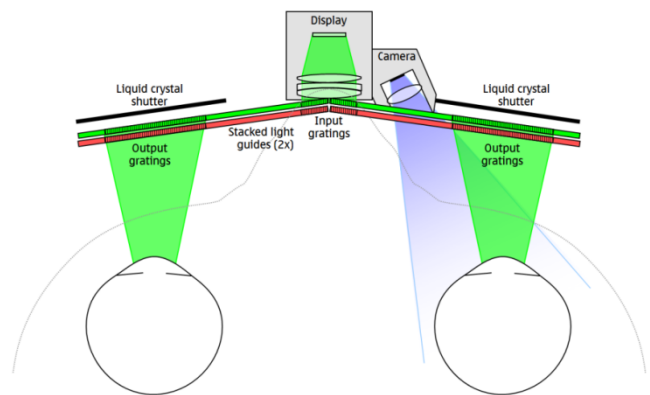


Figure 2. The basic principle of the near-eye display approach with a stack of two light guide plates. The functionality of the

⁴ http://en.wikipedia.org/wiki/Lygia_Clark

⁵ <http://www.distancelab.org/>

⁶ <http://www.distancelab.org/projects/mutsugoto/>

⁷ The code name for this project was Percolate.

light guides is based on diffractive gratings and total internal reflection.

Infrared illumination is arranged for the eye to achieve steady imaging conditions. The eye-looking camera can also be used for tracking the gaze position on the near-eye display virtual image. Gaze is not tracked in this experiment, but the setup affects the nature of the captured images of the eye. Positioning of the camera was originally designed for see-through applications and results in a challenging eye contact in this experiment. For gaze tracking purposes, camera is equipped with a filter which only passes infrared light. This results in monochrome image and possibly slightly unnatural look of the eye (e.g. the color of the iris may seem to be lighter than in reality).

The near-eye display of each user was connected to a laptop PC. The integrated eye-looking camera was connected to the same laptop as an external web camera. The video call was established as a simple Skype call utilizing the eye-looking camera, which also incorporated a microphone, as the video conference camera. Instead of the usual image of the face of the discussion partner in this case only the image of the right eye was visible and displayed in full-screen mode. The system includes a microphone and ear-plug speakers for audio connection. Sound quality in the experiments was standard voice over IP; Skype-to-Skype video calling was used for transmitting both the speech and the eye video. **The quality of the network connection was very high which ensured high audio quality, most likely better than in an average Skype call that participants might have experienced earlier. Still, the audio quality would not be higher than a normal mobile call, so we do not believe that it would alone support the feeling of calmness and intimacy.**

4. THE USER STUDY

4.1. Objectives and Research Approach

To recap the motivation and objectives presented above for this research, they were twofold:

1. From the artistic viewpoint, the aim was to create an image-based communication setup with opportunities for (remote) intimacy between the participants, and explore the potential of this system for evoking feelings and disrupting habitual responses.
2. From the HCI research viewpoint, the objective was to evaluate the user experience (UX) of this kind of novel communication system. **It was not the intention of the system to make emotions visible through a broad communication bandwidth but the more central aim was to support experiences which help people focus and feel tranquility.**

The research approach is constructive – designing and evaluating a novel system – with the aim of understanding dimensions of the UX with this system. Due to the prototypical stage of the system, we were unable to take the system into real contexts of use. Thus, the user experience results are based on empirical data from a user study in a laboratory context.

4.2. Study Process and Methods

The study consisted of five test sessions with one pair of participants in each. The sample size would naturally not be adequate for a full experiment with measurable goals, but our aim was exploration of the designed concept and users' experiences with it, and for that purpose five pairs (ten people) was considered to give adequate feedback.

The sessions started with an introduction to the study and the session, and proceeded with filling out research consents and

background questionnaires. Then each person of the pair were separated into two small rooms where they performed test tasks with the eye-based videoconferencing system. Both rooms had a similar set of the system; the near-eye display showing the other person's eye, and the audio connection.

In the test tasks the pair used the system to discuss with each other about given topics. The topics were chosen in a way that they would include neutral discussion (to start with) but then proceed to more emotional issues. The test tasks were given on paper one at a time to one of the participants who then read each task out loud so that the other one could hear it through the system. The participants were told when to start and stop each task and they were given about 5 minutes for each discussion task. There were five tasks which were as follows:

1. Agree and plan common weekend plans for the next weekend (they can be real or imaginary plans).
2. Talk about the happiest day of your life.
3. Talk about a sad experience you have had in your life over the past year or so. Alternatively, you can talk about a situation which has made you angry.
4. Tell to each other what is going on in your life in the present moment, how are you?
5. Discuss how you experienced this communication session with the near-eye display. How did it make you feel?

The first task was a warm up task for the participants to get used to the system, Tasks 2-4 included topics that could possibly lead to intimate or emotional discussions, and the purpose of Task 5 was to get the participants' first impressions of the experience in a free discussion form.

These tasks were constructed by us based on our best understanding of what could produce emotional content in the discussion. The tasks were used in the pilot test and they seemed to work well in that they brought up some emotional content as well as descriptions of their experiences with the system, so we decided to use them also in the actual test.

After the test tasks, the participants were then interviewed individually. The interview consisted of six questions:

1. How did it feel to communicate with the other person like this and see one of his/her eyes?
2. Did you feel that there was any connection between what you were talking or feeling and the video image of the eye?
3. What was best about the experience?
4. What was worst about the experience?
5. Compare this experience to other communication experiences, a) Skype call and b) regular phone call.
6. Describe your experience with three words.

Each test session lasted for about one hour. The sessions were conducted in English, which was not the native language of the participants – however they all reported they felt confident in speaking English. The reason for this was that the artist who followed the tests was an English speaking person.

Data Gathering. The participants' demographics and background of using different kinds of communication technology were collected with a questionnaire. The whole duration of performing the test tasks and the interviews were video recorded with a camera standing in front of the participant. A laptop computer that

was running the communication application was placed on a table next to the participant and the screen facing the video camera. This setting allowed getting a video recording where we could see the participant and on the laptop screen also the video image that he/she was seeing in the NED glasses (see Figure 3). Audio tracks of the discussions and the video image of the eye were also recorded with the laptops.



Figure 3. Screenshot from the video recording of a participant during test tasks. The eye of his pair is shown on the laptop screen and inside his near-eye display (NED) glasses.

User experience was also measured quantitatively with AttrakDiff questionnaire created by Hassenzahl et al. [8]. AttrakDiff measures pragmatic and hedonic quality aspects of products or systems. We used a version of AttrakDiff that is available in the service provided by the User Interface Design GmbH on their website [3]. The questionnaire is based on a semantic differential with 28 word-pairs in four dimensions: pragmatic quality, hedonic quality – identity, hedonic quality – stimulation, and attractiveness. One of the words in each pair is considered as negative or undesired characteristic of a product and the other word describes a positive and desired characteristic. In our study the participants’ answers were collected using paper forms and afterwards input to the service on the website.

Data Analysis. From the background questionnaire data the participants’ demographics and experience with technology were analyzed with basic quantitative methods. The data of the final user experience questionnaire was analyzed with the AttrakDiff web service, but we also calculated the means and standard deviations of the participants’ evaluations by ourselves for presentation purposes. Video tracks from both persons of each pair were combined and synchronized. The videos of tests tasks 1-4 were viewed and the topics of the discussions were written down. Also other notes were made, for example if the topic clearly affected the participants’ facial expressions. Task 5 and interviews were fully transcribed and analyzed qualitatively by bottom-up thematic coding of users’ statements.

4.3. Participants

Ten participants – five pairs – in total attended the test sessions. The pairs were recruited as a convenience sample from the employees and students of Tampere University of Technology and the employees of Nokia. It was a requirement that the members of a pair were familiar to each other – either in a relationship, work colleagues or friends. It was assumed that with pairs that are familiar to each other there is a greater chance for having intimate or emotional discussions than if the pairs were strangers.

There were five females and five males, all were mixed pairs, and their ages varied between 24 and 41 (mean 33.8, median 33.5). The participants’ demographics are presented in Table 1.

Table 1. Details of the participants.

	Relationship between the pair	How long have known each other	Participant ID	Age	Gender
Pair 1	In a relationship	3.5 years	1A	31	Male
			1B	33	Female
Pair 2	Married	12 years	2A	39	Male
			2B	41	Female
Pair 3	In a relationship	1.5 years	3A	24	Female
			3B	29	Male
Pair 4	Work colleagues	5 years	4A	39	Male
			4B	35	Female
Pair 5	Friends	14 years	5A	33	Female
			5B	34	Male

All of the participants had a university degree on at least master level, most from the field of technology. Nine of them reported having experience of using videoconferencing or internet phone call systems (e.g. Skype) either one-to-one or in a group at least weekly or monthly. Also the one participant (3A) without any first-hand experience of such systems was otherwise familiar with their usage principles. All of them were daily users of computers and mobile phones.

5. RESULTS

The results of the study are presented in three sub-sections. Researchers’ observations of the participating pairs’ discussions related to tasks 1-4 are first summarized briefly (Section 5.1). The experience-related discussions during Task 5 and the interviews were analyzed more thoroughly and coded thematically (Section 5.2). The results of the AttrakDiff questionnaire are presented in Section 5.3.

5.1. General Observations about the Discussion Topics

In Tasks 1-4, the participants talked about a neutral issue (Task 1, their plans for the weekend) and emotionally coloured issues (Task 2: happy experiences, Task 3: sad or angry experiences and Task 4: how they are at the moment).

The participants were quite lively in their discussions, although tasks 2-4 were a little difficult to start for some of the pairs. As happy events, the pairs talked, for example, about memorable trips abroad, their forthcoming wedding, and children’s birthdays. As sad or angry situations, they talked about a disappeared pet, health problems of a family member, and about a conflict with neighbors

or with the boss. The discussion about how they are at the moment was mostly quite neutral talk about the current plans at home or at work.

General observations made by the researchers from the video streams were that the near-eye display (NED) glasses were not totally comfortable for all the participants (especially the ones wearing eyeglasses under the NED); some users tried to use hand gestures to make a point; the laughing and smiling during happy topics could be seen from their eyes quenching and the cheek bones getting higher; but the angry and sad discussions were not as clearly seen from the eyes.

5.2. Results of Participants' Experiences with the System

With regards to participants' experiences of using the system, the following main three themes emerged from the qualitative analysis of data from Task 5 (the discussion about the experience) and the final interview: The eye as a representation of the other person and their emotions, comparison to other means of remote communication and the evaluation of the experience.

It was also analyzed qualitatively if there are clear differences in the opinions of the pairs who are in a relationship and the pairs who are workmates or friends, but no clear differences were found.

5.2.1. The Eye as a Representation of the Other Person and their Emotions

The eye as the visual display. Two of the pairs talked about how their pair's eye does not seem to look straight ahead but a bit to the right. (This was due to the technical fact that the camera inside the glasses was not directly in front of the eye but in the inner corner of the right eye.) One participant said that it is better that way as it does not feel so probing. *"The very fact that the eye is not looking straight at me makes it a bit better. Because it is not so probing."* (2A, Task 5)

Three pairs also discussed how the eye just seemed to be something to look at – they didn't think of it as an eye – but it also reminds that there is a real person. *"...you kind of forget that it is an eye, you just see a black blob."* (2B, Task 5) *"I don't think of it as an eye, I just think it is something to look at. Maybe that's how I know that I'm talking to somebody actually."* (4B, Task 5) Also in the interview two participants said they saw the eye as just something to look at and said they couldn't really connect it to their pair.

One pair pondered if color picture could link the eye more to its owner and *"make it more you"* (5B, Task 5). One participant saw this as a rare experience to *"...stare at someone's eye in such a close range"* (2B, Task 5). Another participant said that even if the eye is so big and staring, it is not scary or uncomfortable. However, she later added that it could become uncomfortable with a stranger, staring at a stranger's eye.

Seeing the eye vs. the whole face. During Task 5 all pairs said something about whether they would prefer seeing only the eye or the whole face. This issue divided the participants' opinions. The ones who preferred the whole face said, for example, that the face would be nicer as the facial expressions can be seen better from the whole face, especially if the mouth is visible. Another explanation was that the whole face is something that people are used to seeing. *"...it would be also nice to see the whole face because that is something that you are also familiar with. Just when I look at your eye I don't know who's there really."* (3B, Task 5)

The ones who liked seeing only the eye said that it is easier to talk to an eye, it makes the discussion more anonymous, or it is enough to see the eye. *"...there is the good thing that you wouldn't see all the face movements and all that, so I can move around and maybe do something else...because only the eye would show. So it makes it a bit more anonymous even though we know each other."* (4A, Task 5)

Focusing the attention on the audio. During Task 5 two of the pairs discussed about the focus of their attention. One of them said that with this system you have to focus on the discussion as you cannot do or see anything else while talking. The other pair also noted that the focus is more on the audio of the discussion than the image of the eye. *"...I wasn't much focusing on what I was seeing, I was more focusing and mostly thinking of what I was hearing."* (2A, Task 5) Also in the interview two participants talked about how it was easier to focus on the audio and the discussion as there was not anything disruptive coming from outside, just the eye to look at.

Identifying emotions from the eye. Identifying emotions from the eye was talked about spontaneously in Task 5 by three of the pairs and it was also discussed with all of the participants in their interviews. In Task 5, one of the pairs said that they think some emotions can be identified by watching the pupil, and also if the person is for example crying. *"You can see what the other person is feeling, somehow. The black part of the eye gets smaller and bigger."* (1B, Task 5)

One pair was stating the opposite, that the emotions cannot be seen from only looking at the eye. They thought that seeing the whole face or at least two eyes would be very helpful to really identify emotions. In addition, one pair discussed that inside the NED glasses it might be too dark to see any changes in the pupil but for example heavy blinking can be seen and it might indicate that the other person is nervous.

Also in the first question of the interview (How did it feel to communicate...?) six participants mentioned identifying emotions and the difference between seeing only one eye or the whole face. Two of them thought it was possible to see some emotions from the eye but it would help to see also the mouth, three participants thought they would need to see the whole face to be able to identify emotions properly, and one merely wondered if seeing the whole face would change the situation. It was also mentioned that some things can be heard from the voice also, for example laughter.

The second question of the interview went deeper into the topic of identifying emotions: Did you feel that there was any connection between what you were talking or feeling and the video image of the eye? Now six participants said it is possible to see something from just the eye but that they didn't notice anything in this particular session. *"The reactions to what I was saying were not necessarily very visible because you only saw the eye, and you didn't see any other parts around the eye which would probably give away feelings a bit better."* (2A, interview, Question 2)

One participant thought the eye and the conversation were separated and there was no connection. Two participants said they noticed something but that it was probably partly because they could also hear things in their pair's voice. *"It is hard to distinguish between the tone of the voice and what you see in the eye, so of course the combination..."* (3B, interview, Q2)

5.2.2. Comparison to Other Means of Remote Communication

Comparison to videoconferencing. Already in the first question of the interview one participant said that his primary feeling was

that this kind of communication is very personal especially compared to an opposite situation of having a conference call in an office cubicle where there are external noises and people passing by. Later in the interview also other participants were asked to compare this eye-based system to traditional videoconferencing (Question 5 of the interview). According to two participants, one of the pros of this system compared to traditional videoconferencing is the aforementioned good quality of the audio. *“Again, the audio aspect of it was very good. There were no disturbances of any kind, it was very pleasant, you could hear well.”* (2B, interview, Q5a)

Two participants noted that emotions are easier to identify when seeing the whole face. Five participants mentioned the fact that with Skype or such systems it is possible to see more of the person you are talking to and the surroundings. *“With Skype you can show things to the camera.”* (1B, interview, Q5a) *“...you can use your hands and show things and it's a bit more lively and this was more intimate, perhaps.”* (1A, interview, Q5a)

Two participants said they feel more comfortable using this system than traditional videoconferencing systems, but had different reason for the opinion: 1) you cannot see the other person's face and know that the other person cannot see your face, and 2) the eye alone seems more honest because there is nothing else to interpret, no gestures or such. One participant said this system might reinforce the feeling of personal and secure discussion, where only two people are present. Two people thought this system helps you to focus on the discussion, as you cannot do anything else while using the system. *“You couldn't get distracted visually to anything. It kind of helped you to focus on the matter at hand. ... It is not as if you could start reading a newspaper, which I sometimes do when I'm using Skype. I'm beginning to wander off to other things.”* (2A, interview, Q5a)

Comparison to a phone call. Four participants said that also compared to phone call this system makes it easier to focus on the conversation as you cannot do anything else at the same time. However, another participant said that especially at the beginning watching the eye took some of her attention away from the discussion. *“Here I needed to focus maybe more and during a regular phone call maybe my thoughts would drift and I would observe what is going around even more and maybe do some other things while discussing. So I found this one better in that sense.”* (4A, interview, Q5b)

Two participants commented seeing the eye: one of them thought this system was comparable to a phone call and the eye did not bring anything new to the experience, whereas the other one thought this system was more personal than a phone call and liked the fact that there is something to see. Two other participants saw some shortcomings in this system compared to a phone call: the glasses were a bit uncomfortable and this system could not be used when walking on a street.

5.2.3. Evaluation of the Experience

Best in the experience. During the interview six participants listed as best about the experience things like the experience being personal, intimate and it felt like being close to the other person. One of them also thought it would be interesting to test the system with someone you don't really know, and see if it would change the feeling. *“It felt a bit like being closer to each other than in normal telephone conversation.”* (1A, interview, q3) *“A strange sense of intimacy.”* (2A, interview, Q3)

Two people said that best was the whole idea because it was so new and interesting to try out. Another one admitted that it was

nice to try but she also noticed that she prefers the traditional kind of communication and would not use this every day.

Four participants mentioned things related to the technical side of the system: the great audio quality enhanced the feeling of being close to the other person, everything worked well, and it was comfortable. Also during Task 5 the great audio quality raised discussion among three pairs and one participant even said *“...it is like as if you were here in this room with me”* (2B, Task 5). Two people said they liked this system because the other person does not see their face, but only the eye, which makes it feel less exposing. *“It was easier to talk to an eye than to just see the other person and to actually know that the other person sees my whole face.”* (4A, interview, Q3) One person described the experience as *“pleasantly strange”* (2A, interview, Q3).

Worst in the experience. All pairs talked briefly about the equipment during Task 5 and mentioned mostly negative issues. The NED glasses were not fitting very well on some of the participants, especially those who kept their eye glasses under them, and one pair wondered if they would start feeling dizzy or their eyes would get tired if they used the system for a longer time. Also during the interview most of the participants, eight in total, mentioned technical issues when they were asked about worst things in the experience: their eyes got tired, the NED glasses were heavy and uncomfortable, the system was not the most natural one to use, the display could be better quality, and the other person's eye did not look straight at them. *“I guess the technical things that... It was so heavy.”* (1B, interview, Q4)

One participant said during Task 5 that the display seemed fine otherwise but he could not see properly the small picture of his own eye which was shown in the bottom right corner of the view. Two participants said in the interview that they would rather have wanted to see the whole face of their pair than only one eye. One participant added that one of the worst issues for him was also that the system did not seem to bring any new value to communication.

The experience as a whole. Four pairs mentioned in their discussion on Task 5 that the experience was interesting. It was also seen as new, unordinary and surprising. One pair said it felt tranquil as there were no outside noises. It was also seen as more personal than using just the voice as in a normal phone call.

Also in question 1 of the interview five participants described the experience as interesting. Three participants used words like unusual, odd and weird, but in a positive sense. *“It was interesting; it was new, something that I hadn't experienced before.”* (4A, interview, q1) *“It was pleasant, very pleasant in an unusual way. One that didn't really seem very ordinary.”* (2A, interview, Q1)

In Task 5 one participant said he didn't see the point of this kind of system at all. Relating to this, another one said it was like watching a film, except that there wasn't much going on. *“But I have to say that I somehow failed to see the point. Especially when I kind of hear you smile but I don't see it. It doesn't show up in the eyes and the anger doesn't show up in the eyes.”* (5B, Task 5)

Words used to describe the experience. In the answers to question 6 of the interview the experience was described with 23 different words or expressions which are illustrated in Figure 4 in alphabetical order. Some of the participants could think of only two words and some of the words were mentioned by several participants. The font size indicates the words' popularity and the most popular word was “interesting” with three mentions and “different”, “fun”, “intimate”, “new”, and “surprising” were each

mentioned two times. The rest of the words in were all mentioned once.

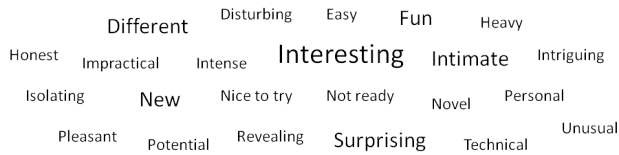


Figure 4. Participants' words describing their experience of eye-based videoconferencing.

5.3. Results of the AttrakDiff Questionnaire

The AttrakDiff user experience questionnaire [3] included 28 word-pairs on a 7-point semantic differential scale from -3 to 3. The participants selected a spot on the scale that described their experience of communication with the eye-based videoconferencing system. Figure 5 presents the means of the participants' answers, with the standard deviation bars.

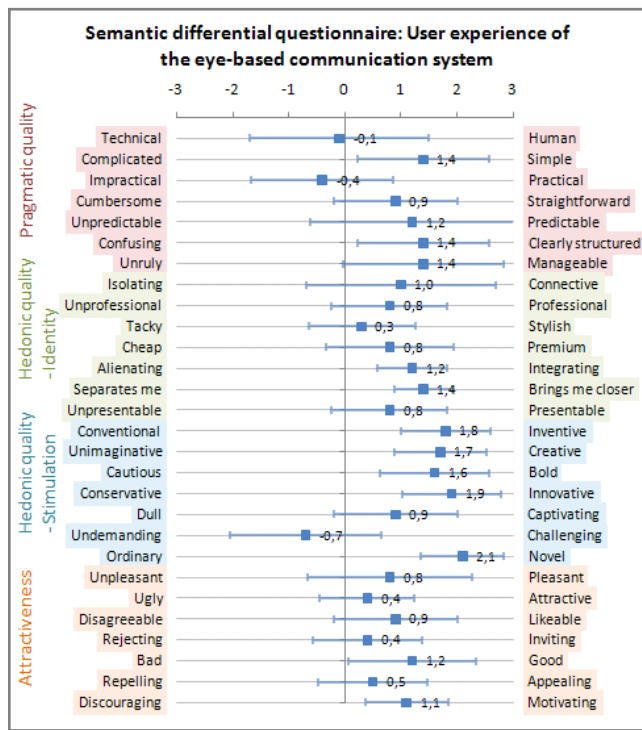


Figure 5. Results of the AttrakDiff questionnaire.

From Figure 5 it can be seen that communicating with the eye-based system was experienced as positive in most aspects that are included in the word-pairs. The most positive means are related to the dimension of hedonic quality - stimulation and more precisely to the word-pairs ordinary-novel (mean: 2.1), conservative-innovative (mean: 1.9) and conventional-inventive (mean: 1.8). Some scores, especially on attractiveness, are near neutral. This may be at least partly due to the technical nature of the test set-up. The mean scores of the evaluations are negative on three of the word-pairs. Two of them are related to the system's pragmatic quality: the system was experienced slightly technical (mean: -0.1) and impractical (mean: -0.4). In addition one negative mean value related to hedonic quality (stimulation): the system was experienced more undemanding than challenging (mean: -0.7). However, in the context of the eye-based communicating aiming

at intimacy, the system being described as undemanding could also be seen as a positive quality.

6. DISCUSSION, CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a system for person-to-person videoconferencing with which the users see one eye of the other person in full-screen view on near-eye displays. The system was motivated by an artist's desire to help people 'bare their souls'. Thus it was never intended to be a practical system for efficient business communication. The aim of the user study was to understand how this artistically motivated system would be experienced by the users. More specifically, we wanted to understand if, and how, this system would support the feeling of intimacy between the users, and if the users would have experiences that were "out of the ordinary" and which would help them seize the moment – *carpe diem*. To this end, we conducted an explorative user study in the laboratory context with five pairs of people who knew each other beforehand. The results lead to some interesting findings.

The participants had very variable reactions to whether the communication partner's eye can convey emotions and whether it is a pleasurable "view" to the other person. For many of the participants, the eye was not necessarily a concrete eye, but a "blob" to focus on. The eye can "force" the persons to focus on the issue under discussion – it prevents distractions. Furthermore, audio became important when the visual distractions were brought to such minimal stage. Some users were still missing some of the cues that are usually present in nonverbal communication. Many participants stated that they would prefer seeing the whole face, to gain more understanding of the other person's emotional reactions. Still, the eye alone may feel more "honest" because there is nothing else to interpret.

The fact that the eye did not look directly into the other person – due to the technical setup of the system – might have had some effect on (lack of) feeling the connection with the communication partner. However, many participants felt it was easier to talk to just the eye; it was less intrusive than the whole face.

When compared to other means of remote communication (video conferencing and phone calls), communication with this system can be experienced personal and secure. There are no disturbances, and thus one has to focus on the one-to-one discussion. In the beginning, the eye may take attention away from the discussion because it is an unusual view during communication. However, the participants got used to it and the focus shifted to the discussion at hand.

The overall experience of using the system was reported by the participants as personal, intimate and supporting the feeling of being close to the other person. The AttrakDiff questionnaire results underline the positive experiences of *stimulation* which is related to supporting the "out of the ordinary" experiences – which were one of the main goals of the system. Overall, the AttrakDiff results indicate the more positive (than negative) evaluation of the system.

It may feel less exposing when one's own face is not shown. Only one person stated that they did not see the point of the system at all. The most used terms to describe the system were *interesting*, *different*, *intimate*, *fun*, *new* and *surprising*. Whereas some of these expressions could be seen as neutral, users furthermore described the experience as "a strange sense of intimacy", "pleasantly strange" and "tranquil". These terms resonate well with the original artistic motivations of the system: taking the people out of their ordinary conversation situations and habitual

responses, and to create intimate communication. The eye did not reveal too much of the communication partner's emotional state – a broader view of facial gestures would be needed to interpret those emotional cues. Still, the eye seemed to function as a means to prevent distractions from the focus of the discussions, and thus helped users to stay in the moment.

In terms of the study set-up, it is clear that a laboratory study does not fully represent the actual usage. Due to the prototypical stage of the system we were, however, unable to take the system into real contexts of use. However, we believe that the fact that all the couples knew each other well helped in creating a feel of communicating “for real”. Further studies are needed in real contexts of use, in people's real lives to confirm that the system does indeed support intimacy in long-term usage.

The fact that the participants were using English – a non-native language to them – as the task language might have an effect to how they experience the communication. It is possible that using a non-native language might make discussions less emotional and perhaps, the system usage less pleasurable. However, we believe that the fact that all pairs knew each other well (and had voluntarily come to the test with each other) reduced the threshold of discussing topics with emotional content. Naturally there is some limitation to the depth of the discussions that can take place in this kind of test, as they were recorded, but this can only be overcome by a long-term test in real contexts of use.

Another issue is that the fact that the participants were familiar with technological innovations may have made them more positive to this kind of novel communication approach. However, we believe that we can get more rich feedback when people are feeling relatively comfortable with trying new types of systems.

Conclusions. The main contribution of the study was that the unusual presentation of a single eye of the communication partner can support focus and experience of intimacy. While it is well known that facial expression, gestures and bodily language all play roles, in this study the point was about isolating the eye, away from all of the other signals people read, to attempt to create more focus and intimacy. It was not the intention of the system to make emotions visible with the broadest possible multimedia bandwidth but the more central aim was to help people focus and feel tranquility with a limited form of visual communication.

We feel that overall, the aims of enhancing a feeling of intimacy between conversants via disrupting usual modes of communication was achieved by the system. Also, the idea was to create more of a focus on the moment, (could we point back to the [attraktdiff here?](#) _ and the subjects taking part in the study reported feeling more focus on the conversation. After considering the results, we wonder if this idea of ‘focus’ could provide the distinction between hearing and listening. Listening happens when understanding is involved, when our minds give meaning to the information. Thus, when true listening occurs, a greater connection and a feeling of intimacy can arise. Therefore, the study made us relook at the concept, and we are thinking of reframing it as a *tool for deeper listening*.

Also, subjects reported usage of the eye-based videoconferencing system being *tranquil*. We originally thought that the intervention (using the system) would have created more of a feeling of vulnerability than tranquility. The artist had originally conceptualized the intervention as a tool for ‘baring your soul’ and had presupposed that the feeling of vulnerability would create a feeling of openness and connection. When the artist originally discussed the idea with people, most people looked squeamish at the idea of staring into one's eye and the sense of vulnerability it may ignite. However, during participation people found it

strangely relaxing. In this ‘relaxation’, the eye seemed to get transposed into something both intimate, something to focus on, and something to read emotion from. Yet people also reported that the eye became ‘abstracted’. Maybe this abstraction made the eye feel ambiguous, allowing more room to evoke a more meditational feeling of connection and emotion? Poetically, it makes us wonder about the eastern meditation idea of ‘The third eye’. They see this ‘third eye’ a gateway to one's inner world, to yourself. Being more ‘in yourself’ could be a way of being centered and focused in your day to day activities. This may be one way towards a happier life.

Future work. The system could be developed further to support the direct gaze into other person's eye. Also, we could explore a version of the system where both eyes are shown. To this end, a prototype of near-eye display setup with two eye-looking cameras (one for each eye) located on the visual axes could be built. This would allow more realistic transmission of the eye-based videos, and thus, possibly, a deeper emotional connection. Alternatively, mouth, or multiple parts of the body could be explored. Also, instead of the black-and white displays in the presented setup, it would be interesting to explore the interaction with eye-based displays using colour.

It would also be valuable to take the system out of the lab, to be used in people's real lives and everyday contexts. It would also be exciting to undertake a study where the participants did not know each other. A long-term study of the system's effects to personal communication could be conducted to explore the system's potential to help people calm down and enjoy the moment in their daily interaction with other people.

7.ACKNOWLEDGMENTS

[Removed for blind review.]

8.REFERENCES

1. Aaltonen, V., Takatalo, J., Hakkinen, J., Lehtonen, M., Nyman, G. and Schrader, M. 2009. Measuring Mediated Communication Experience. International Workshop on Quality of Multimedia Experience, QoMEx 2009. Vol., no., pp.104-109, 29-31 July 2009. DOI: 10.1109/QOMEX.2009.5246967
2. Argyle, M. and Dean, J. 1965. Eye-Contact, Distance and Affiliation. *Sociometry*, Vol. 28, No. 3, pp. 289-304.
3. AttrakDiff, a service of User Interface Design GmbH, <http://www.attrakdiff.de/en/Home/>. Cited 1.12.2010.
4. Cacioppo, J. and Patrick, W. 2008. *Loneliness: Human Nature and the Need for Social Connection*, W. W. Norton & Company.
5. Cha, J., Eid, M., Rahal, L. and El Saddik, A. 2008. HugMe: An Interpersonal Haptic Communication System. *Haptic Audio visual Environments and Games, 2008. HAVE 2008. IEEE International Workshop on*, vol., no., pp.99-102, 18-19 Oct. 2008. DOI: 10.1109/HAVE.2008.4685306
6. Grayson, D.M., and Monk, A.F. 2003. Are You Looking at Me? Eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction*, 10, 3 (September 2003), 221-243. DOI=10.1145/937549.937552
7. Hassenzahl, M. 2003. The Thing and I: Understanding Relationship Between User and Product. In: Blythe, M.A., Overbeeke, K., Monk, A.F., and Wright, P.C. (Eds.) *Funology: From Usability to Enjoyment*. Kluwer, pp. 31-42.

8. Hassenzahl, M., Burmester, M. and Koller, F. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In J.Ziegler & G. Szwillus (Eds.), *Mensch & Computer 2003. Interaktion in Bewegung* (pp. 187-196). Stuttgart, Leipzig: B.G. Teubner.
9. Hauber, J., Regenbrecht, H., Hills, A., Cockburn, A. and Billinghurst, M. 2005. Social Presence in Two- and Three-dimensional Videoconferencing. In *Proceedings of the 8th Annual International Workshop on Presence*, London, September 21-23, 2005.
10. Heikkinen J., Olsson T. and Väänänen-Vainio-Mattila K. 2009. Expectations for User Experience in Haptic Communication with Mobile Devices. *Proc. of MobileHCI'09*, ACM.
11. IJsselsteijn, W., van Baren, J. and van Lanen, F. 2003. Staying in Touch: Social Presence and Connectedness through Synchronous and Asynchronous Communication Media. In Stephanidis, C. and Jacko, J., editors, *Human-Computer Interaction: Theory and Practice (Part II)*, Volume 2 of the *Proceedings of HCI International 2003*, pages 924–928. Lawrence Erlbaum, Hillsdale, NJ.
12. Ishii, H., Kobayashi, M. and Grudin, J. 1993. Integration of Interpersonal Space and Shared Workspace: clearboard design and experiments. *ACM Trans. Inf. Syst.* 11, 349–375.
13. REMOVE FOR BLIND REVIEW Järvenpää, T. and Äyräs, P. 2010. *Highly Integrated Near-to-eye Display and Gaze Tracker*. *SPIE Photonics Europe, IEEE*.
14. McLuhan, M. 1994. *Understanding the Media: The Extensions of Man*, MIT Press.
15. Rosenthal, A.H. 1947. *Two-way Television Communication Unit*. United States.
16. Vertegaal, R., Slagter, R., Van der Veer, G.C. and Nijholt, A. 2001. Eye Gaze Patterns in Conversations: There is More to Conversational Agents Than Meets the Eyes. Paper presented at the *ACM CHI 2001 Conference on Human Factors in Computing Systems*, Seattle.
17. Werner, J., Wettach, R. and Hornecker, E. 2008. United-pulse: Feeling Your Partner's Pulse. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services (MobileHCI '08)*, ACM.