Multitasking Behavior and Gaze-Following Technology for Workplace Video-Conferencing

Eveline Van Everdingen

RESEARCH MASTER ARTIFICIAL INTELLIGENCE – COGNITIVE SCIENCE
VRIJE UNIVERSITEIT AMSTERDAM

OCTOBER 30, 2015

TABLE OF CONTENTS

PREFACE

The idea of studying multitasking behavior in business video-conferencing has been developed at the beginning of my internship, at the research laboratory of Fuji Xerox in Palo Alto, California[1] (FXPAL). I have set up an experiment to investigate what factors are influencing perceptions of, and attitudes towards multitasking behavior in video-conferencing. This work is described in a paper[2], added to the thesis in Appendix A. Different types of an experimental design can be applicable researching this topic. Though, consistency between subjects and conditions is an important factor in this subjective field of research. Therefore, we decided to record different types of common technologies used for multitasking activities and ask participants' opinions about the movie. This paradigm showed such successful results, that the same experimental design has been employed to search for improvements regarding perception of multitasking behavior in a follow-up study. The follow-up study, researching how dual-monitor multitasking can be perceived more positively, is one of the future work implications evolving out of the first study. I conducted this study (described part 1 of the thesis) at the VU after my internship at FXPAL. The study gives a theoretical basis for the application of a two-webcam gaze-following technology to keep the frontal view of the multitasker and sense of eye-contact. The sense of eye-contact is disrupted when the video-conferencing partner is focusing on the second monitor; their gaze falls outside the scope of the webcam mounted on the first monitor. I have developed a simple video-conferencing setup for dual-monitor workplace settings (technical work described in part 2). Based on facial feature detection, I designed a webcam selection mechanism combining two webcams, each camera mounted on a monitor. The systems assumes that the user is looking at the monitor of which the corresponding webcam has detected a frontal face. In this case, the other person will always be able to observe the face of the multitasker, instead of half a face or cheek. Preliminary analysis of an explorative implementation show promising results which can be used for future implementations.

Each part of this thesis can be read independent from each other. To have a better idea of how multitasking is being perceived, it is advised to start with Appendix A. This study is conducted in cooperation with Dr. Marlow and Dr. Avrahami at FXPAL. Results of this study are described in a paper and submitted to the conference CSCW'16. The first part of this thesis investigates how dual-monitor multitasking can be perceived more positively. The developed two-webcam selection tool is investigated and used for the experiments in this study. The second part describes the technical implementation of the developed mockup of the gaze-following system based on facial feature detection. I have performed and written the follow-up study independently, after my internship at FXPAL. Though, I have been collaborating with FXPAL and combined my work with their work for a second paper on this topic. This paper is submitted for publication at the conference IUI'16[3]. My work in this paper represents a short version of the application study on the two-webcam video-conferencing setup (part 1) and a description of the two-webcam selection system based on facial feature detection (part 2). All experimental materials including experimental movies, dataset, analysis and technical code, related to part 1 and part 2 of the hesis are included in Appendix B and can be accessed at: [http://dx.doi.org/10.6084/m9.figshare.1619896].

---

[1] http://www.fxpal.com/

[2] Marlow, J., van Everdingen, E., & Avrahami, D. (2016). Taking notes or playing games? Understanding multitasking in video communication. To appear in *Proceedings of CSCW 2016.*

[3] Avrahami, D., van Everdingen, E. & Marlow, J. (2016). Supporting multitasking on video-conferencing using gaze tracking and on-screen activity detection. Submitted to *Proceedings of IUI 2016.*

Part 1. How to support the multitasker? A study on gaze-following video-conferencing technology

This part investigates how dual-monitor multitasking behavior can be perceived less negatively. This is based on prior research presented in Appendix A. The data of this study has been used for describing the evaluation section of gaze-following technology in a paper[4] co-authored by me which is currently under review.

[4] Avrahami, D., van Everdingen, E. & Marlow, J. (2016). Supporting multitasking on video-conferencing using gaze tracking and on-screen activity detection. Submitted to *Proceedings of IUI 2016.*

How to support the multitasker? A study on gaze-following video-conferencing technology

Eveline van Everdingen
Vrije Universiteit Amsterdam
e.a2.van.everdingen@student.vu.nl

ABSTRACT

Dual-monitor multitasking during a video-conference is not perceived as positive compared to single-monitor multitasking. Dual-monitor multitasking might disrupt feelings of eye-contact and engagement with the remote partner, which are important factors for positive evaluation of multitasking behavior. This study investigates a gaze-following video-conferencing design, hypothesizing that improving the sense of eye-contact will result in a better video-conferencing experience with the multitasker. Perceptions of the multitasker's behavior are judged as more polite and acceptable in gaze-following technology, although the sense of eye-contact is not significantly more positive rated with this experimental design. These results show that gaze-following webcam technology can be successful to improve collaboration in dual-monitor multitasking.

INTRODUCTION

Nowadays, working on multiple monitors is very common. It is known to work more efficient and to bring structure as well as a better overview in a job [23]. Even in business video-conferences, dual monitors are used. Although the purpose of dual monitor use might be clear to the multitasker, this behavior is not always perceived as positive by their video-conferencing partners. Previous results (Appendix A; [16]) show that multitasking on a dual screen or mobile device is indicated as less polite and acceptable than doing something else on the same screen. Multitasking on a dual screen was indicated as more obvious than on a single screen. Although the multitasker might be involved in the meeting, he or she seems less engaged with the meeting, resulting in negative perceptions.

One of the main differences between these workplace setups is the position of the head relative to the webcam. Switching gaze to another window at the same monitor will only affect gaze fixation. Switching attention to a second monitor or mobile device requires head movements and in turn changes the observer's view of the multitasker, losing the sense of eye-contact. Eye-contact is an important factor for effective communication and functions to provide information, regulate interaction, express intimacy and exercise social control [13]. Although the standard desktop-webcam does not enable direct eye-contact, prior work has shown that people are able to have a certain feeling of eye-contact during video conferencing [2], [8]. Losing the sense of eye-contact can be one of the underlying factors influencing the perceptions of multitasking behavior.

Whether video-conferencing technology can improve the perceptions of multitasking behavior, is investigated in this study. Many different types of technology can be researched in more detail to encounter negative effects of switching between devices in a video-conference. As described in the section 'design implications' of the first study, new designs may be developed to *increase transparency of what the other is doing*, or to *help the multitasker be perceived more positively*. As proposed, a two-webcam video conferencing setup following gaze might address (some) negative effects when switching between monitors. This video-conferencing setup might give the observer a better understanding of what the multitasker is doing and maintain the sense of eye-contact. Investigating these effects, the goal of this research will be to answer the following questions:

How can gaze-following video-conferencing technology positively affect perceptions of multitasking behavior?

What role does the feeling of having eye-contact play?

Answering these questions will lead to a better understanding of how multitasking behavior is being perceived and affects spatially diverted communication and collaboration. This knowledge contributes to design as well as technical development of computer supported cooperative work. Multitaskers using multiple monitors will benefit being more aware of how their multitasking activity is being perceived and what technology will result in a better representation of their behavior.

The structure of this paper is as follows: the importance of perceiving eye-contact in (video-mediated) communication will be explored first. Then, a short overview of the work done on gaze-following video-conferencing technology follows. To investigate the effect of gaze-following video-conferencing technology, a behavioral survey is conducted investigating perceptions of eye-contact and multitasking behavior.

Eye-contact in face-to-face communication

In face-to-face meetings, failure to maintain eye contact increases feelings of mistrust [2]. The level of perceived eye contact, or what someone has been told about the partner's level of making eye contact, affects the perception of eye-contact with the partner [12].

Gender plays an important role in eye-contact and perceptions of the partner's personality. In general, females tend to have more eye-contact in a dyadic conversation than males [15]. Avoidance of gaze in men has been related to being "emotionally inhibited, over-controlled, and to having psychosomatic and physical symptoms". Gaze avoidance in women is associated with a higher degree of "psychopathy, hysteria and traditional feminity" [14].

In a study by Kleinke *et al.* (1973), participants were asked to rate their and their partner's level of searching for eye contact after a 10 minute male-female conversation. When both, males and females, were told to have received more eye contact than average, they rated their partner as being more sincere. Partners were rated as less attentive when participants were told to have received less than average gaze. When females were being told that their male partners had an average gaze, they rated their partners as less relaxed, whereas males rated their female partners as more relaxed [12]. These results suggest that the (believed) *feeling* of having eye contact is an important factor in communication and perceived attitude towards the other.

Memory can also be affected by the level of eye-contact. More products were recalled when presented by a prerecorded salesman pitch when the salesman was making eye contact with the camera, than when he avoided eye contact [7].

These studies indicate that in face-to-face communication, the level of perceived eye-contact influences how we perceive our remote partner and what we remember from the conversation. People are able to have a certain sense of eye-contact even if there is no ability to have direct eye-contact. The next section will describe previous research on the sensitivity of perceiving eye-contact and when people lose this feeling in video-mediated communication.

Losing the sense of eye-contact

The most common individual-to-individual video-conferencing setup is a webcam placed on top of the monitor or laptop. However, this setup hinders the ability to have direct eye-contact with the other person. When both video-conferencing partners are looking at each other's eyes on the monitor, the webcam will capture a downward gaze direction. To capture direct gaze fixation, gaze should be focused direct into the webcam, as if the other is making eye-contact. However, people are more intended to look at their screen, observing their partner's face, than looking into the black whole of the camera. Much research has been done on actual eye-contact and whether people do have the the feeling of eye-contact, or lose this feeling of eye-contact in video-mediated communication.

The range of gaze directions that gives someone the feeling to be looked at, is called the 'cone of gaze' [8]. The cone of gaze has been researched in reality and virtual environments. Direction of gaze can best be described as cone-shaped, with a varying visual angle of 0 (center of the cone) up to a maximum between 4 and 9 degrees (edge of the gaze cone). The maximum angle depends on the position and distance of the looker. Since direct eye-contact is measured as the center of the cone (0 degrees) and people still indicate eye-contact up to 9 degrees gaze deviation, it can be concluded that people are able to have a feeling of eye-contact even though there is no actual direct eye-contact. The bigger the distance between the observer and the looker, the wider the 'cone of gaze' angle. Therefore, a bigger distance increases the likelihood that the observer will have the feeling to be looked at [8]. Though, sensitivity to deviations is not similar in each direct of the cone of gaze. Horizontal gaze deviations are slightly more accurate than vertical directed deviations [4], [6].

The feeling of eye-contact can also be explained from a more neuropsychological perspective. Boyarksa *et al.* (2015) investigated the 'Mona Lisa Effect', the feeling of being followed when changing position. The authors studied brain patterns of participants when viewing full face pictures with a varying angle of the eyes. Similar brain activation was found when viewing the slightly averted face (cone of gaze was 10 degrees) as well as the face on the edge of the feeling of eye-contact (cone of gaze was 5 degrees). This indicates that similar signals are processed when someone has the feeling to be looked at and slightly more averted from that.

The mentioned studies, as well as other studies have researched the ability of having a certain sense of eye-contact, with varying results in the exact angles of the cone of gaze [6], [8], [10]. This is not a surprising result, since there are many methodological differences between the studies. Overall, it can be concluded that in direct and video-mediated communication: 1) detecting actual eye contact is harder than having the feeling of eye-contact, 2) people are very sensitive for deviations in gaze, 3) although a bit less sensitive to downward gaze, and 4) head position, orientation and distance to the camera/ the other influences the perceptions and sensitivity of gaze direction. These assumptions indicate that even in video-mediated communication people can perceive eye-contact with the other, even though mutual eye-contact is disrupted by the standard video-conferencing technology.

Effect of gaze perception

Having the feeling of eye-contact with the other in video-mediated communication positively affects trust. Bekkering *et al.* (2006) investigated the effect of the viewing angle of the speaker in a video message on ratings of trust. Video messages were recorded with different viewing angles of the communicator, frontal view (with direct eye-contact), side-view and upper-view (without direct eye contact). Participants had lower ratings of (the feeling of) eye-contact, when the speaker was captured from the side or above, as expected. These speakers had also lower ratings of trust, compared to the speaker of the frontal-view captured message. Other work showed that camera positioned on top of the monitor makes the person look taller and the person is being perceived as more influential than when the webcam is placed below [11].

More positive effects of eye gaze have been found. Increased levels of eye-contact is related to higher ratings of the others engagement with the conversation [17]. Prior work even indicated that people like each other more when they are able to observe the others face and gaze [1].

In short, eye-contact and observation of the frontal face are positively affecting ratings of trust, engagement with the conversation and perceptions of the other. These results are strong indicators for positive effects when making use of a two-webcam gaze-following technology when dual-monitor multitasking is a necessity. Therefore, some of these indicators are used to investigate the effect of the proposed gaze-following setup for video-conferencing.

RELATED WORK

Both hardware and software solutions have been invented to overcome the webcam-monitor eye-contact discrepancy. The video tunnel might be one of the earliest solutions to this parallax. A video tunnel is a combination of a semi-transparent mirror, placed in an angle of 45 degrees relative to the user, as well as a screen placed behind the mirror, in front of the user. The user is able to look through the mirror to look at the image of the remote partner. The camera is placed in 90 degrees of the user, capturing the mirrored image. Vertegaal *et al.* (2002) used this setup for multi-party video conferencing meetings with multiple webcams for each participant in the meeting. An eye-tracking system detects which camera the speaker is facing so that the others will be able to have a frontal-face view of the speaker. Switching between cameras resulted in an angular shift in the transmitted image of the speaker. However, the angular shift did not strongly affect eye-contact perception and was not considered to be distractive [25].

Instead of using a video tunnel to overcome the camera parallax, other setups has been developed placing cameras behind the display. MAJIC enables eye-contact by placing a camera behind a transmissible screen on which the view of the interlocutor is projected [20]. Positive effects were found with this type of video-conferencing setup, with increased ratings of perceived trust [22].

MultiView video-conferencing technology focuses on improving perceptions of gaze and other non-verbal cues in multiple spatially diverted groups [18]. MultiView is a video-conferencing setup combining multiple webcams, in which each webcam corresponds to one participant. To each individual a unique perspective of the other group is presented as if the groups are talking face-to-face. Instead of one webcam capturing the whole group, different angles are captured corresponding to the positions of the individuals observing the others, and vice versa. Participants indicated to have a sense of eye-contact during the entire video-conference. In a follow-up study, measures of trust between groups have been investigated. Groups who used the MultiView video-conferencing setup, presenting each individual the corresponding viewing angle of the other group, had significantly higher ratings on trust compared to one universal webcam view presentation of the other group [19]. These results indicate that a certain sense of eye-contact positively contributes to collaboration in video-mediated meetings.

Next to development of hardware technology, other work related to software artificially adjusting gaze direction improved perceptions of eye-contact in video-mediated communication. Computer vision techniques have been used to detect and adjust the focus of the eyes and position of the head. Adjustments are made in such way that it looks like someone is facing the camera. After analysis of each individual video frame, the eye contour, head orientation and gaze direction will be determined. A new picture showing the eyes in the desired direction is built upon a virtual 3D model of the head [9].

There have been invented multiple customized solutions of hardware and software implementations for eye gaze correction [21]. However, the most prevalent technology is the standard desktop with a webcam mounted on the monitor nowadays. This setup is the most cost-efficient, portable and easiest way to use for individual-to-individual video-mediated communication [2]. Though, the standard one-webcam monitor setup shows its shortcomings in a dual-monitor workplace. Working on a dual monitor during the video-conference is seen is impolite and unacceptable, compared to single-screen multitasking. Perceptions of eye-contact are disrupted when switching to another monitor. In this study we investigate whether a two-webcam gaze-following setup can capture the multitasker and maintain the sense of eye-contact with the multitasker, making the multitasker be perceived less negatively.

EXPERIMENTS

Hypotheses

In a classic video-conferencing setup, the camera parallax obstructs the ability to have direct eye-contact with the other. However, previous research indicates that with this video-conferencing setup, people can perceive *the feeling* of having eye contact. In turn, the feeling of having eye contact positively affects communication and collaboration as previously described. Therefore, it can be theorized that when someone is multitasking during the video call, maintaining the feeling of having eye-contact can be in favor of the perceived attitude towards multitasking behavior. Leading to the first hypothesis:

H_1: Having the feeling of eye-contact will positively affect the attitude towards multitasking behavior.

Dual monitors are mostly used to search for extra information or to make a physical deviation in tasks or projects per monitor. It improves efficiency and structure of the work environment, resulting in higher overall performance. When the user switches focus to another monitor or area in the room,

Figure 1. Webcam view of person B engaged with the meeting (left), multitasking on second monitor with switching to another webcams (center) and multitasking without switching to another webcam (right) during the video-conference.

the interlocutor will observe a side view of the multitaskers' face and will lose (the feeling of) eye contact. Therefore, the technical set-up of the video-conference (VC) determines the view of the video-conferencing partner and might influence the remote partner's perceptions of eye-contact.

H_2: The VC setup will have a significant effect on the perceived eye-contact with the multitasker.

Whereas in a single-webcam video-conferencing setup, someone will lose the feeling of having eye-contact with the multitasker when the attention is switched to a second monitor.

H_{2a}: In a single webcam VC setup, switching to another monitor during the VC will reduce the sense of eye-contact with the multitasker.

However, when a frontal face view of the multitasker is captured during the multitasking activities, eye-contact might be regained. Another webcam placed on top of the second monitor should capture the multitasker when focusing on the second screen.

H_{2b}: In a dual webcam VC setup, switching to another monitor during the VC will not reduce the sense of eye-contact with the multitasker.

As previously described, (the feeling of) having eye-contact is an important factor for communication and collaboration. Losing perceptions of eye-contact as a result of the multitasking activities might negatively affect the attitude towards multitasking activities. When the video-conferencing setup is able to maintain or prevent perceptions of eye-contact, this factor might also affect the perceived attitude towards multitasking behavior.

H_3: The VC setup will have a significant effect on the perceived attitude towards the multitasking behavior.

Method

To examine these hypotheses, participants are asked to watch a 1-minute movie about two people having a work-related video call. The same method for this experiment has been used as in [16] (see Appendix A). One person (person A) describes three different potential locations for their advertising campaign. The other person (person B) acts passive, responding with 'yeah', 'okay' or 'alright' to person A. After a notification sound, person B gets distracted with something else and switches to another monitor, multitasking.

To test the effect of the different video-conferencing setups while multitasking on a second monitor, two different webcams placed on top of each monitor have been combined and recorded the acting of person B. The first camera is located on top of the primary monitor, which is the monitor person B is facing at the beginning of the conversation. Another webcam is located on top of the dual monitor, which is used for multitasking activities during the conversation. When the multitasker (person B) switches gaze to the dual monitor, the proper webcam will start capturing the multitasker. The technical implementation of this system is described in the second part of this thesis [part 2]. Recordings of person A have been reused from study [16] (see appendix A) for this experiment.

Out of this one take of recording op person B, two different movies have been constructed. The first movie is constructed out of the recordings of the primary webcam only (*Static View* condition). For the second movie, the recordings of the 'second' webcam are added to the primary webcam recordings, when the multitasker switches to the dual monitor. When person B switches gaze, the movie will, simultaneously with person B, replace the viewpoint from the primary webcam to the 'second' webcam. Then, after repositioning of person B to the primary monitor, the view of this person follows gaze and changes to the primary webcam (*Dynamic View* condition). In figure 1, the view of person B is illustrated when engaged with the meeting, together with both views when person B is multitasking on a dual monitor (*Static vs. Dynamic View*). Another movie was constructed in which no multitasking acitivities were performed to check whether this behavior has been rated significantly different from the multitasking conditions. Person B does not switch between monitors and is recorded with one webcam during the meeting (*Control group).*

In study [16] (see Appendix A), two different types of commonly used interfaces were included. Next to the 'side-by-side' view of equally large frames of person A and B, was the 'big-and-small' view included in which the multitasker was presented in the larger frame. When the multitasker was

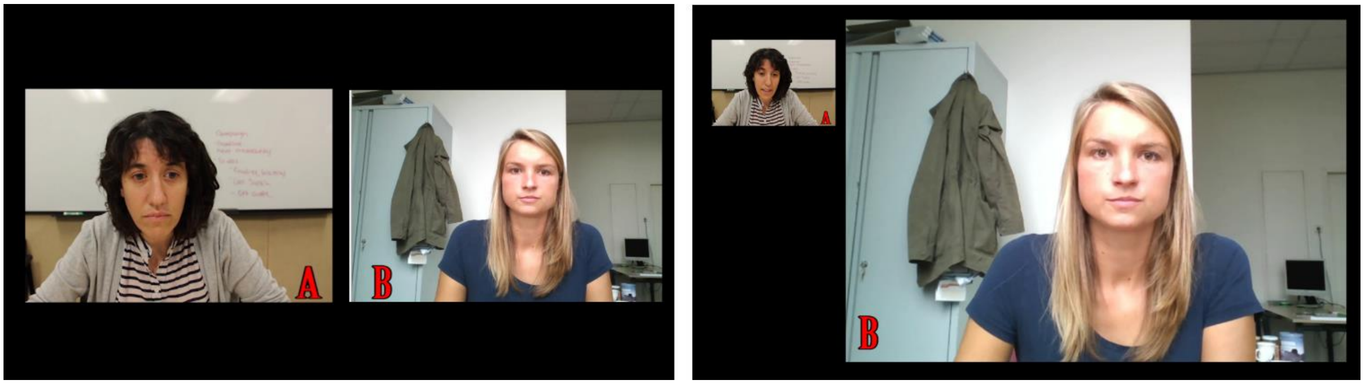

Figure 2. Interface layout with equal frame size, side-by-side (left) and the multitasker presented in a visually larger format (right).

presented in a visually larger frame, behavior of the multitasker was rated significantly different. One explanation for this could be that participants are paying more attention to person B, compared to person A. The video conferencing setting might also feel more natural and the participant is more actively involved in the online meeting.

Trying to create an experimental design which approaches the natural environment even better, the three conditions described (*Static View* vs. *Dyncamic View* of Webcam Technology and a *Control* group) are tested in two experiments. For each experiment, the mimicked video-conference is designed in different visual layouts. In the first experiment, frame sizes of person A and B are equally large, placed side-by-side. In the second experiment, which was conducted after the first one, person B is presented in a visually larger format compared to person A, see fig. 2. Except for the two different types of interface layout of the video conference are the experiments exactly the same.

In total, the experimental design contains three conditions; two types of Webcam Technology (*Static View*, *Dynamic View)* and one group without multitasking *(Control group)*. This design is tested in two experiments. In the first experiment the Interface Layout of the video-conference movie is in a 'side-by-side' view, in the second experiment the movie is formatted in a 'big-and-small' view. (See Appendix B for the movies)

Measures

This study is a follow up on study [16] (see Appendix A), therefore most of the measurements are reused. Participants are asked to answer open-ended questions about what they saw person A and B doing after watching the movie first. Then, they have to answer which locations have been discussed in the advertising campaign. These answers give a better idea of what participants were thinking and whether they paid attention to the movie.

Similar questions about the attitude towards multitasking behavior (H_1 and H_3) have been used as in study [16]. The statements are rated on a 5 point Likert-scale ranging from "Completely disagree" to "Completely agree". The following statements are asked to rate:

1. Person B's behavior was *polite* during the conversation
2. Person B's behavior was *acceptable* during the conversation
3. It was obvious to me that Person B was multitasking
4. It was obvious to Person A that Person B was multitasking
5. Person A was engaged with the meeting
6. Person B was engaged with the meeting

To test H_2, participants are asked to rate on the same 5-point Likert scale their Perceived Eye Contact with person B during the conversation *("I had the feeling of having eye-contact with person B most of the time.").*

Subsequently, to have a better idea about the obviousness of switching between webcams, participants are asked to give an open-ended answer about the obviousness of switching between webcams (*"Please explain in a few sentences how obvious it was to you that person B was recorded with different webcams during the conversation.")* and they are asked to indicate on a scale of 1-6 how many times the view of the webcam was changed during the VC.

At last, demographics of each participant is included in the survey; age, gender and their experience with video-conferencing tools (such as Skype, Google Hangouts etc.).

Procedure

Participants are online recruited with the crowdsourcing website CrowdFlower[5]. Only individuals located in the United States have been accepted for the study. A link on the CrowdFlower website redirects the participants to one of the three randomly assigned online surveys (conditions). The survey takes about 4 minutes and after finishing the survey, participants receive a unique code which has to be entered in the CrowdFlower page to earn their reward of 40 cents (to make sure that participants will complete the survey).

[5] http://www.crowdflower.com/

The participants are asked to watch a 1-minute movie, which has a fixed page-timer of 1 minute preventing the participants to skip the movie. In addition to the fixed page-timer, another hidden timer also records the time needed for each page of the online survey. These timers are used for detecting outliers; people whose completion time was too short or too long might not have taken the survey seriously.

Participants

Individuals are only allowed to participate in the study once (between-subjects design) and different participants are recruited for the experiments. Before analyzing the data to investigate the proposed hypothesis, the surveys are screened for bad respondents who did not pay attention. This is very important when using an online platform for participant recruitment, as there is not much control on quality of the responses. Based on IP address, double takes are removed. Low quality of work is detected with: 1) the location check question, 2) a similar response check of Likert-scale questions, 3) wrong indication of switches between webcams (only in the *Dynamic View* condition) and 4) Mahalonobis Outlier Analysis on the completion time.

Experiment I

In total, 259 participants responded to the survey in the first experiment, of which 161 respondents participated more than once and 26 surveys were detected on low quality of work. After cleaning the dataset, 72 participants were kept for further analysis with an average of 24 participants per condition (Min=23, Max=25 and SD=.82). 60% of the participants were men.

Experiment II

In the second experiment, 333 participants have responded, of which 96 responses are removed for multiple participation and 85 unreliable respondents are filtered on low quality of work. In total, the dataset of the second experiment consists of 160 participants with an average of 53 participants per condition (Min=48, Max=63 and SD=6.85). 61% of the participants were men.

RESULTS

Experiment I

The effects of multiple webcams used for dual-monitor multitasking is investigated with an analysis of variances (ANOVA) of Webcam Technology on the measures of Politeness, Acceptability, Engagement of person B, Obviousness of Multitasking and Perceptions of Eye-Contact. Age and Gender are included as covariates. In this experiment, no significant effects are found for the type of Webcam Technology (*Static View* vs. *Dynamic View)* used on each of the dependent measures.

To investigate whether the same effect is found for the different types of Interface Layout on the perceived engagement with the meeting of the multitasker, an analysis of variance (one-way ANOVA) is conducted combining both datasets of the experiments (excluding control groups). The type of Interface Layout is included as independent measure

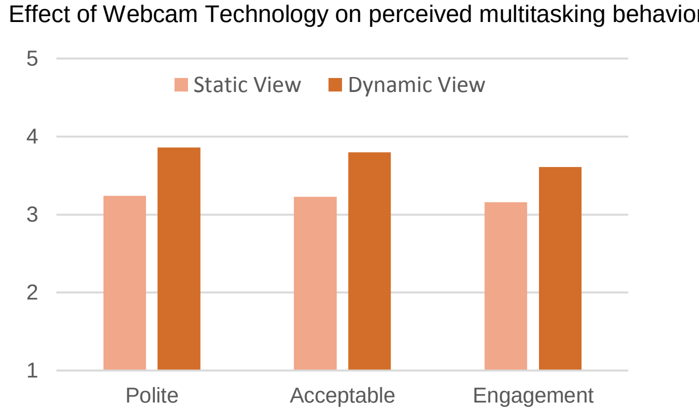

Figure 3. Ratings of Politeness, Acceptability and Engagement are significantly higher in the Dynamic View condition with a two-webcam setup.

and Engagement of person B as dependent measure. Indeed, similar results are found as in study [16] (see Appendix A). Interface Layout has a significant main effect on the perceived engagement of person B ($F[1,156]=6.20$, $p=.014$). A post-hoc analysis student's t-test showed that participants significantly rated the multitasker (person B) to be more engaged with the meeting when presented in a visually larger format (M=3.4) than in similar format with person A (M=2.8, $t(156)=2.49$, $p=.014$).

Experiment II

Effect of Webcam Technology used for video-conferencing

H_3 (the type of Webcam Technology will have a significant effect on the perceived attitude towards multitasking behavior) is tested with an analysis of variances (ANOVA) of Webcam Technology as independent measure and ratings on Politeness, Acceptability and Engagement of person B as dependent measures. Age and Gender are included as covariates, the control group is excluded in the analysis.

Results show a significant main effect for Webcam Technology on the ratings of Politeness ($F[1,107]=9.05$, $p=.003$) and Acceptability ($F[1,107]=7.64$, $p=.007$). Multitasking behavior observed in the *Dynamic View* is rated significantly higher on Politeness (M=3.9) and Acceptability (M=3.8) than in the *Static View* (M=3.2 for both measurements). Therefore, H_3 can be confirmed. Analysis of the effect of the multitasker's engagement with the meeting ($F[1,107]=4.61$, $p=.034$) shows similar results. The multitasker is perceived as more engaged with the meeting when presented in the *Dynamic View* (M=3.6) then in the *Static View* (M=3.2). These results are illustrated in figure 3.

Effect of Perceived Eye-Contact

Whether the type of Webcam Technology will have a significant effect on the Perceptions of Eye-Contact, H_2, is investigated with an analysis of variances (ANOVA) of Webcam Technology as independent measure and ratings on Perceived Eye-Contact as dependent measures. Age and Gender are included as covariates.

No significant main effect are found for the type of Webcam Technology used on the Perceived level of Eye-Contact (F[1,107]=2.42, *p*=.123). Although the *Dynamic View* was rated higher than the *Static View* (M=3.4 vs. M=3.1) on average, these differences are not significant. Therefore, H_2 and both sub-hypotheses, H_{2a} and H_{2b}, must be rejected.

H_1 stated that the feeling of eye contact will have a positive effect on the attitude towards multitasking behavior. A Spearman's rank-order correlation of Perceived Eye-Contact and Politeness, Acceptability and Engagement of the multitasker has been conducted to test this hypothesis. Results show a positive correlation between ratings of Perceived Eye-Contact and Politeness (r=.47, p<.001), Acceptability (r=.54, p<.001) and Engagement of person B (r=.48, p<.001). These results indicate that higher perceived eye-contact with the multitasker correlates with more positive perceived behavior of the multitasker. Therefore, H_1 can be confirmed.

Obviousness of switching between webcams

Participants were asked to describe the activity of person B (the multitasker). In the *Static View* condition, more negative descriptions of person B's behavior are given than in the *Dynamic View* condition. For example, in the *Static View* participants wrote; *"Person B seemed distracted and looked at someone who was in the room"* and "*paying attention but also seemed to be checking emails"* and "S*he was listening, but was distracted".* Multitasking activities described in the *Dynamic View* are mostly related to working on another monitor, some examples; *"Person B was listening, and then looked into a different monitor before switching back"* and *"Closely watching her webcam...on her two monitors"*, even more positive descriptions of person B were given; *'She was listening carefully and paying attention to person a and she was looking at the email that person a has sent.'*.

In the *Dynamic Switch* condition, people are also asked to give an open-ended answer on how obvious it was that the view of the multitasker switched between the two webcams. The answers are analyzed and coded into 'obvious' or 'not obvious', neglecting other answers. 38 participants indicated that switching was 'obvious', whereas 6 participants indicated 'not obvious'.

Although switching viewpoints was very obvious and multitasking was not hidden, open ended-answers about the multitasking activity are more positive and more related to the meeting in the *Dynamic View* than in the *Static View* condition. This is in favor of the proposed gaze-following webcam technology when multitasking during video-conferencing.

Effect of Interface Layout

Finally, the two control groups of both datasets (the different types of Interface Layout of experiment 1 and 2) are compared to investigate whether the level of eye-contact is influenced by the representation of the multitasker in the video-conference.

A one-way analysis is conducted with Interface Layout as independent measure and ratings of Perceived Eye-Contact as dependent measure, age and gender are included as covariates. The main effect of Interface Layout has a significant effect on the Perceived Eye-Contact (F[1,70]=3.38, *p*=.05). Post-hoc analysis shows significantly higher rates of Perceived level of Eye-Contact in the second experiment with the 'big-and-small' view than in the 'side-by-side' view of the first experiment (M=3.8 vs. M=3.2, *t*(70)=37.37, *p*=.043).

No main or interaction effect of Interface Layout and Webcam Technology is found on the perceived Eye-Contact or Politeness and Acceptability of multitasking behavior.

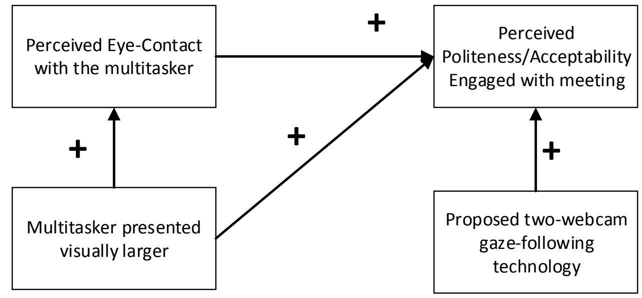

Figure 4. Two-webcam gaze-following technology and perceived eye-contact positively affects attitudes towards dual-monitor multitasking behavior.

DISCUSSION

The goal of this study was to investigate how multitasking on a dual screen can be perceived as less negative and what role eye-contact plays. Two settings of Webcam Technology have been compared for multitasking on a dual monitor while video-conferencing; 1) standard desktop-webcam setup with one webcam mounted on one monitor and 2) the proposed gaze-following setup with a webcam mounted on each monitor. Figure 4 shows the effects found in the experiments investigating the proposed two-webcam video-conferencing setup.

Multitasking on a dual monitor is not perceived as polite and acceptable when gaze falls outside the scope of the webcam. However, when another webcam is taking over the frontal face view when the multitasker switches to another webcam, the attitude towards this behavior is more positive. People rate the multitasking behavior as more polite and acceptable and the multitasker is even perceived as more engaged with the meeting. These results cannot be addressed to a certain 'hiding effect' of multitasking. Participants indicated that switching between the two webcams was very obvious as well as it was related to multitasking. However, the feedback of the participants about person B's behavior suggests that activities of the multitasker are perceived as meeting-related rather than something else. Switching between webcams also gives an indication of the presence of a second monitor, which can be attributed to meeting-related activities such as searching for useful information. We can conclude that the observation of the frontal face is an important factor for a pleasant video conference and that the proposed two-webcam gaze-following technology positively contributes to perceptions of the multitasker.

It was expected that the underlying mechanism for this effect was partially due to the ability to keep the sense of eye-contact when the multitasker switches to another monitor. Contrary our expectations, ratings of the feeling of having eye-contact with the multitasker did not differ between participants observing the multitasker with one webcam (losing the frontal face view when the multitasker switches to the second monitor) or gaze-following webcam technology (keeping the frontal face view of the multitasker).

Multiple causes can be addressed to this lack of increased feelings of eye-contact. The most plausible cause might be that the design of this study is performed by a third party, where the design forces them to observe two people having a conversation, instead of actively participating in the conversation. Even though this 'third party experimental design' has its benefits, where the scenario can be kept as similar as possible between subjects, it also requires a certain level of empathy of the participant to become fully engaged with the meeting. Another cause explaining the contradicting results can be the total time the multitasker is focused on the second monitor. In total, the multitasker is engaged with something on the dual screen 25% of the time (16 sec. of 1:05 min.). Stretching the time being engaged on the second monitor might have a bigger impact on the observers' perceptions of the multitasking behavior.

The level of perceived eye-contact is not directly related to the type of video conferencing technology used. However, higher ratings of perceived eye-contact is positively related with the judgments of the multitasker's behavior. The more participants were having the feeling of eye-contact, the more positive the multitasking behavior was judged. This indicates that when a video conferencing system facilitates the user to have (the feeling of) eye-contact when multitasking, their behavior will be perceived as more positive, in favor of the proposed gaze-following system.

In line with the study [16] (see Appendix A) the visual representation of the multitasker in the video conferencing application did play a role on how multitasking behavior was judged. Showing the multitasker in a visually larger format compared to the video-conferencing partner affects the way the multitasker is perceived. Representation is also affecting perceptions of eye-contact, which is stronger when the multitasker is presented in a bigger format than the other person. Although the sense of eye-contact was not improved when using the proposed two-webcam system, people are more convinced that the multitasker in engaged with the meeting. Even perceptions of multitasking behavior are more positive, one of the most important effects of this system.

Limitations and future work

One limitation of this study is that subjects participated in a passive way observing the video-conference, as earlier described. An implication for future work investigating communication in video-conferencing technology can be setting up an experiment in which participant have to interact with the multitasker in a video-mediated conversation. For example, participants have to collaborate in a task in which they have interact with someone else (the 'multitasker') to achieve a goal. For this task, both persons will have the ability to work on a dual-monitor. The other person (the 'multitasker') can be intentionally involved with something related or unrelated to the task on the second monitor. This empiric research creates a more natural setting than the one used in the current studies [part 1 and 2]. Participants might feel more engaged with the video-conference, with their opinions more truly perceived rather than opinions based on an artificial video conversation.

Another limitation and implication for future work is that this study only focuses on improving perceptions of dual monitor multitasking. Although this type is very common at the workplace, mobile devices and tablets are also one of the main technologies used for multitasking. More research can be done on improving perceptions of eye-contact by combining other technologies to multitask with in the video-conferencing software. For example, when interaction with a mobile device is detected, the camera built in the smartphone can be turned on to show the multitasker. As proposed in study [16] (see Appendix A), other ways of showing what the multitasker is doing can be done by feeding messages or alerts of the interrupting technology to the video-conferencing screen. Both parties will benefit, the 'message-receiver' is able to read the message on the screen (without switching gaze off the screen), whereas the observer of the multitasker (the 'message-receiver') has a better understanding of what the other is doing.

CONCLUSION

This study gives strong indications for positive effects of two-webcam gaze-following video-conferencing technology on attitudes towards dual-monitor multitasking. First, the sense of eye-contact is positively influences attitudes towards multitasking behavior. Second, continuously presenting the frontal face view of the multitasker makes the multitasker look more engaged with the online meeting. At last, a visually larger representation will positively affect the sense of eye-contact. These results support the proposed two-webcam gaze-following tool for video-conferencing. Future work should investigate how other types of technology used for multitasking can be perceived as more positive, improving the sense of eye-contact with the multitasker and transparency of secondary activities.

REFERENCES

1. Bailenson, J., Beall, A. & Blascovich, J. (2002). Gaze and task-performance in shared virtual environments. *Journal of Visualization and Computer Animation*, *13*, 313-320.

2. Bekkering, E. & Shim, J.P. (2006) i2i Trust in video conferencing. *Communications of the ACM, 49,* 103-107.

3. Bohannon, L. (2010). Effects of video-conferencing on gaze behavior communication. *Thesis/Dissertation Collections. Rochester Institute of Technology.* Extracted from: http://scholarworks.rit.edu/cgi/viewcontent.cgi?article=2340&context=theses [Accessed at September 2015].

4. Bohannon, L., Herbert, A., Pelz, J., & Rantanen, E. (2013). Eye contact and video-mediated communication: A review. *Displays, 34,* 177-185.

5. Boyarskaya, E., Sebastian, A., Bauermann, T., Hecht, H. & Tuscher, O. (2015). The Mona-Lisa effect: Neural correlates of centered and off-centered gaze. *Human Brain Mapping, 36,* 619-632.

6. Chen, M. (2002). Leveraging the asymmetric sensitivity of eye contact for videoconference, in: L. Terveen (Ed.), *Proceedings of CHI'02,* ACM, 49–56.

7. Fullwood, C. & Doherty-Sneddon, G. (2006). Effect of gazing at the camera during a video link on recall. *Applied Ergonomics, 37,* 167-175.

8. Gamer, M. & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 705-715.

9. Gemmell, J., Toyama, K., Zitnick, C. L., Kang, T. & Seitz, S. (2000). Gaze-awareness for video-conferencing: A software approach. *Proceedings of CSCW'00,* ACM, 25-35.

10. Grayson, D. & Monk, A. (2003). Are you looking at me? Eye contact and desktop video conferencing. *ACM Transaction on Human-Computer Interaction, 10,* 221-243.

11. Huang, W., Olson, J. & Olson, G. (2005). Social dynamics can be distorted in video-mediated communication. *Doctoral Dissertation, University of Michigan, Ann Harbor.* Extracted from: http://dl.acm.org/citation.cfm?id=1087837&coll=DL&dl=GUIDE [Accessed at September 2015].

12. Kleinke, C. Bustos, A., Meeker, F. & Staneski, R. (1973). Effects of self-attributed and other attributed gaze on interpersonal evaluations between males a females. *Journal of Experimental Social Psychology, 9,* 154–163.

13. Kleinke, C. (1986). Gaze and eye-contact: A research review. *Psychological Bulletin, 100,* 78-100.

14. Larson, R. & Schakelford, T. (1996). Gaze avoidance: Personality and Social judgments of people who avoid direct face to face contact. *Personality and Individual Differences,* 21, 907-917.

15. Levene, M., Sutton-Smith, B. (1973). Effects of age, sex and task on visual behavior during dyadic interaction. *Developmental Psychology, 9,* 400-405.

16. Marlow, J., van Everdingen, E. & Avrahami, D. (2016). Taking notes or playing games? Understanding multitasking behavior in video communication. To appear in *Proceedings of CSCW'16,* ACM, xx-xx.

17. Nakano, Y. & Ishii, R. (2010). Estimating user's engagement from eye-gaze behaviors in human agent conversations. *Proceedings of IUI'10,* ACM, *139–148.*

18. Nguyen, D. & Canny, J. (2005). MultiView: Spatially faithful group video conferencing. *Proceedings of the CHI'05,* ACM, 799-808.

19. Nguyen, D. & Canny, J. (2007). MultiView: Improving trust in group video conferencing through spatial faithfulness. *Proceedings of CHI'07,* ACM, 1465-1474.

20. Okada, K., Maeda, F., Ichikawaa, Y. & Matshushita, Y. (1994). Multiparty videoconferencing at virtual social distance: MAJIC Design. *Proceedings of CSCW'94,* ACM, 385-393.

21. Polycom EagleEyeDirector. Extracted from http://www.polycom.com/products-services/hd-telepresence-video-conferencing/realpresence-accessories/realpresence-accessories-eagle-eye-director.html [Accessed at September 2015].

22. Regenbrecht, H., Müller, L., Hoermann, S., Langlotz, T. Wagner, M. & Billinghurst. M. (2014). Eye-to-eye contact for life-sized videoconferencing. *Proceedings of the Computer-Human Interaction Conference on Designing Futures: the Future of Design*, 145-148.

23. Ross, S. (2003). Two Screens Are Better than One. *Microsoft Research.* Extracted from: http://research.microsoft.com/en-us/news/features/vibe.aspx [Accessed at October 2015].

24. Vertegaal, R., Weevers, I. & Sohn, C. (2002). GAZE-2: An attentive video conferencing system. *Proceedings of CHI Extended Abstracts '02,* ACM, 736-737.

25. Vertegaal, R., Weevers, I.. Sohn, C. & Cheung, C. (2003). GAZE-2: Conveying eye contact in group video conferencing using eye-controlled camera direction. *Proceedings of CHI'03,* ACM, 521-528.

Part 2. Dynamic two-webcam gaze-following technology based on frontal face detection

In this part, a technical implementation of gaze-following technology with two webcams is described. The effect of this technology is described in part 1. Each webcam is placed on one monitor and selection of the proper webcam is based on frontal face detection. The work done on this implementation is included in a paper[6] currently under review, co-authored by me.

[6] Avrahami, D., van Everdingen, E. & Marlow, J. (2016). Supporting multitasking on video-conferencing using gaze tracking and on-screen activity detection. Submitted to *Proceedings of IUI 2016.*

Dynamic two-webcam gaze-following technology based on frontal face detection

Eveline van Everdingen
Vrije Universiteit Amsterdam
e.a2.van.everdingen@student.vu.nl

ABSTRACT
Multitasking on a dual-monitor can increase work-efficiency, it is also often rated as impolite and unacceptable. Having the feeling of eye-contact in a video-conference positively affects the perceptions of multitasking behavior and communication. However, when multitasking on another monitor, the perceptions of eye-contact are disrupted. In this paper, an explorative system will be presented that enables the view of the multitasker's frontal face to keep the feeling of eye-contact with the multitasker. Computer vision is applied for frontal face detection to select the webcam which presents the optimal view of the multitasker's face. The current implementation shows promising results for future gaze-following technology supporting dual-monitor video-conferencing.

INTRODUCTION
In case of using multiple monitors, the standard desktop-webcam video conferencing system does not fulfill the user's needs. Observation of the multitasker's frontal face is impossible when switching gaze to another monitor, resulting in negative perceptions of the multitasker [part 1]. Capturing the multitasker with another webcam placed on top of the second monitor might cover these shortcomings, promising results have been found [part 2]. As previously proposed, this system should simultaneously to the user's gaze direction switch capturing the multitasker with the corresponding webcam. Based on frontal face tracking, the system should select the proper camera. In this case, the observer will always be able to look at the multitaskers' frontal face in de video-conference.

This paper describes a preliminary implementation of an automatic webcam selection tool, based on facial feature detection. Prior work related to face detection in video conferencing tools is described first. Then, the constraints of which an ideal system should comply with is summarized, followed by the actual implementation. After this, an evaluation describes a rough indication of the performance of the system and improvements for future work.

RELATED WORK
In previous telecommunication tools, simple eye-tracking hardware has been applied to determine which direction the user is gazing [6]. Later, the ability to acquire, analyze and understand images with computer vision software has enabled feature detection and recognition of the human face [2].

Not much video conferencing tools have been designed with face detection as a camera selection mechanism. However, one commercial tool has been developed already. Perch[7] telepresence portal connects two iOS devices through face-detection. Perch is an always-on video conferencing App, which gets activated and starts recording audio when a familiar face is detected and recognized. This tool enables hands-free videoconferencing, the user only has to face the webcam. Multiple parties can be connected together and form a group conversation.

More expensive solutions are developed for multi-group video-mediated meetings. For example, Polycom[8] designed a multi-webcam video conferencing tool which recognizes the speaker and selects the closest webcam to capture the speaker in a conference room with multiple participants [4]. Many video conferencing tools have been developed for multiple group collaboration, resulting in highly expensive customized business telecommunication systems. However, not much research has been done to simple individual-to-individual video-mediated meetings working on dual-monitors. This paper will describe an implementation of a simple two-webcam selection mechanism based on face detection for video-mediated dyads.

DESIGN CONSTRAINTS
In order to design a proper video conferencing tool, we defined the following criteria, based on previous research [part 1 and 2];

1. Assuming that one of the multiple monitors is used as the primary screen to work on and another monitor for secondary/other activities, the corresponding primary webcam should get activated when starting the video conference.
2. When the user switches to another monitor, the proper webcam should have positive results for frontal face detection and should start recording.

[7] https://perch.co/

[8] http://www.polycom.com/

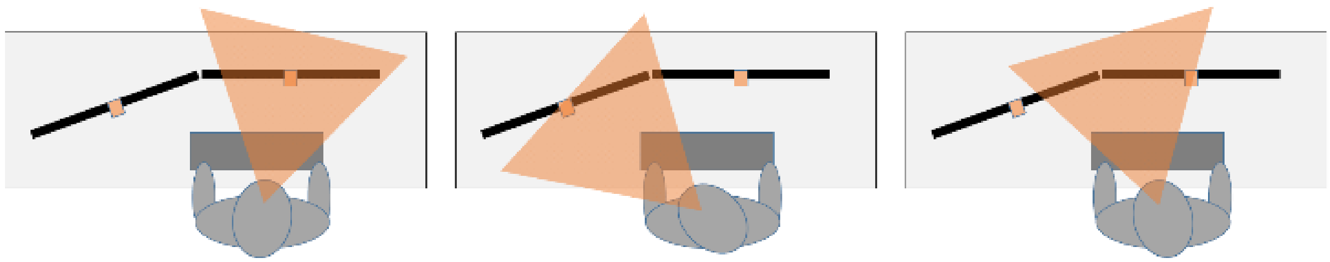

Figure 1. Gaze direction of the multitasker with the focus on the primary monitor (left), on the dual monitor (middle) or in between two monitors when switching (right).

3. The switching between these two webcams should have a 'natural' feeling. Meaning that when the head of the user is moving towards the second monitor, the second webcam should start recording at the moment the gaze direction of the multitasker is in between the monitors. In this way, the other observes the multitasker gazing away from the observer first, followed by the video stream showing the multitasker heading towards the camera (captured by the second webcam). A certain optimum in the sensitivity of switching should be found, too fast switching between cameras can be confusing, but switching too slow can be very annoying and does not contribute to the goal of the system (capturing the users' frontal face).
4. When switching between monitors, the user's gaze is in between the two monitors at some point. Both webcams can have positive results on frontal face detection (or even none of the webcams detect the users face). In this case, the system has uncertain about deciding which camera is actually the best one. Therefore, the system should ideally be able to compute the gaze direction of the user. For each webcam the angle of the users gaze relative to the camera should be computed. The smallest angle indicates which monitor the user is actually facing. Switching between monitors is illustrated in figure 1.

The following section describes how these constraints are approached in the implementation.

IMPLEMENTATION

The code is implemented with Python[9] with the Open Source Computer Vision[10] (OpenCV) library for face selection. This code is available at:
https://github.com/een450/MasterProject.

Face detection algorithm

The OpenCV package comes with pre-trained classifiers for multiple objects, including; frontal face, eyes, mouth, upper body, full body and even eyes with glasses. The object detection algorithm is based on the initially proposed Haar-feature algorithm combining multiple classifiers in a cascade training structure [5].

The classifier is trained with a dataset of positive and negative images (images with and without faces). Three types of Haar-like features, in line with the convolution matrix, are used for object detection use in computer vision; Edge Features, Line Features and Four-Rectangle Features (see fig. 2). Each feature has the value of the difference between two opposing pixel areas.

To detect these features in a high-quality image would take endlessly long. Therefore, a cascade combination of multiple classifiers significantly reduces the total image screening time, without losing object detection quality. When the first classifier has a positive feature detection, the next evaluation will be triggered. When there is a negative evaluation outcome, the evaluation process for the sub-window ends immediately. This method reduces the total amount of negative evaluations considerably, increasing the classification speed. The training classifiers are constructed with AdaBoost [1]. After training, the threshold is minimized to reduce false negative evaluations.

In total, three types of decisions play a role in order to reduce the total amount of evaluations; 1) the number of classifier stages, 2) the number of features in each stage and 3) the threshold of each stage. In any case, there is a trade-off in number of features used for classification. Adding more

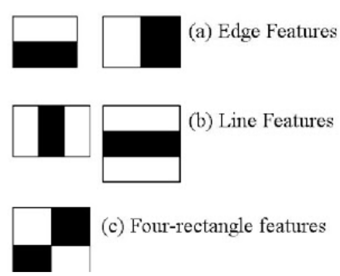

Figure 2. Viola-Jones Haar Features, extracted from [3].

[9] https://www.python.org/

[10] http://opencv.org/

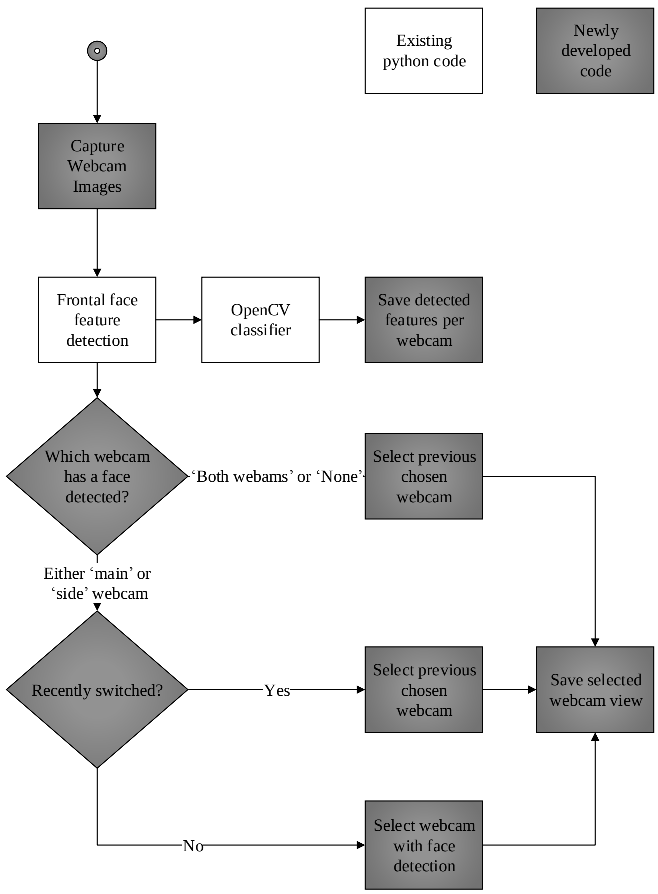

Figure 3. Implemented decision model of the two-webcam system. The system selects the camera which shows the multitasker best in a frontal view to the interlocutor.

features will likely achieve higher detection rates and lower false positive rates, but it also increases computation time.

In short, the HaarCascade classifier is a fast pattern matching algorithm in a series of 'waterfalls', cascades, breaking the whole face detection process into smaller pieces, slightly more detailed in each step. This classifier is known for its short computational time. This is useful for the proposed system which should be able to analyze multiple frames per second of the webcam video feed.

Webcam-selection decision model

The system starts capturing the user with the webcam of the primary monitor, as determined in criteria 1 of the design section. Then, an image frame is captured analyzed in a cycle for both webcams (with a certain pace of frames per second). In each cycle, the system is keeping track of which camera has been decided to present its image in the video and whether the camera system is the same as the previous selected camera or not (switching or not switching between cameras). The decision model on which webcam should be selected to capture the user is illustrated in figure 3.

After the acquisition of images from both cameras, the images are classified at hand of the face detection algorithm with the OpenCV. If classification shows positive results, a list of coordinates is returned outlining an area in which a frontal face view is detected. Ideally, not only the detected frontal face area is found, also the direction of gaze can be determined. However, the latter is not included in the first implementation of the automatic camera selection system due to time management. The system does not meet criteria 4.

Based on the results of the classification process, one webcam image is selected to show in the video conference. When a face has been detected on *both* webcam feeds or *no face* has been detected at all, the same camera decision is made as in the previous cycle to show its current image. This is a substitute-solution for criteria 4. Since only one image can be shown in the video conference, the system decides which image will be presented. Selecting the previously chosen camera reduces (unnecessary) switching between webcams, part of criteria 3.

When the classification for *one* of the two webcam-images shows positive results, the system determines if switching between webcams has occurred recently. *'Recently switched'* is determined at hand of the list of saved previous chosen webcam-views. If the list consists of similar previously selected webcam-views, it can be concluded that the system did not recently switch. In this case, the system is allowed to show the image in which a face is currently detected. Then, the system has permission to switch webcams if the other webcam has detected frontal face view. If the list of 'previous selected views' does not contain similar webcam-views, the system is not allowed to switch and has to select the camera-view similar to the previous cycle. This part is the core of the decision model, determined in criteria 2.

Sensitivity of switching is implemented in the total range of the list containing the previously picked webcam-view. The longer the range, the more similar webcam-views should be gathered to get permission to switch. The shorter the range, the faster the system is allowed to switch. Depending on the processor and frames-per-minute captured by the webcam, the range of the list can be determined. The adjustable range of similar previously selected webcam view is the implementation of criteria 3.

EVALUATION

To get an approximation of how the system performs, a small evaluation is conducted. The system is tested at three different workplace locations, each with proper daylight, two standard webcams each mounted on a monitor. The system has been tested by the author, switching ten times from the primary monitor to the dual monitor and back. Similar workplace setup has been used as illustrated in figure 1. The distance of the webcam mounted on the primary monitor to the user is approximately 40-50cm and the distance between the 'dual-monitor' webcam and the user is app. 50-60 cm. The monitors are positioned in a natural setting, in an angle of 125 º, such that the webcams are at a distance of 50 cm of each other.

On average, 9 out of 10 times that the user switched gaze to the dual-monitor or back, the system follows gaze by switching webcam-views within 1 second. The other one time, the system eventually switched in more than 1 second. When the focus of the user was fixed on one monitor, the matching webcam-view was continuously shown. There were no interrupting changes of camera-view during gaze fixation. These results indicate a stable dynamic gaze-following webcam tool within the boundaries of the test-conditions.

FUTURE WORK

Although this evaluation shows promising results, a more explorative investigation of the system shows certain shortcomings of the current system. These shortcomings are indications for future work.

First, the current technology is best functioning when only one person is captured in the frame by the webcam. For example, people walking by at the workplace will be captured in the background of the image of the user. Therefore, more research should be done to background noise reduction. Similarly, the system has a lower performance when the user is wearing glasses or when other attributes are captured by the webcam. Frontal face detection is not very consistent with facial features such as eyes and nose. With glasses or other attributes, face detection is worse.

The decision which monitor the user is facing, is only based on frontal face detection in the current system. Future research should focus on implementation of user behavior and interaction with the monitor. User behavior, such as an estimation of gaze direction based on position of the eyes and nose relative to the face, will give more information about which area the user is gazing. With these measurements, switching between monitors can already be anticipated when the user's head is moving. Interaction with the monitor, such as mouse location and clicks, are likely to be gazed at. More user feedback will give a better impression on what the user is actually doing. A better understanding of the activities of the user will lead to a better estimation of gaze direction with more precise performance of gaze-following technology.

The implemented system is not a full working video-conferencing system yet. This mockup only combines two webcams and decides which view should be selected based on frontal face detection. The next step is to implement this feature in a video-conferencing software, such as Skype[11] or Google Hangouts[12].

CONCLUSION

A mockup of an automatic two-webcam selection mechanism is developed based frontal face detection. This system is an explorative study showing promising results for future gaze-following webcam technology for individual-to-individual video-conferencing. Results of this explorative implementation are promising for future work. More research should be done to the implementation of user behavior and technology interaction, for a better approximation of user's gaze-direction. Then, this feature can be implemented in existing video-conferencing software.

REFERENCE

1. Freund, Y. & Shapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer Science and System Sciences, 55,* 119-139.
2. Gemmell, J., Toyama, K., Zitnick, C. L., Kang, T. & Seitz, S. (2000). Gaze-awareness for video-conferencing: A software approach. *CSCW, 2000.*
3. Mordvintsev, A. & K, A. (2013). OpenCV-Python Tutorials Documented. Extracted from: https://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html#face-detection [Accessed at July 2015]
4. Polycom EagleEye Director, Datasheet. Extracted from http://www.polycom.com/content/dam/polycom/common/documents/data-sheets/eagleeye-director-ds-enus.pdf. [Accessed at October 2015]
5. Viola, P. & Jones, V. (2001). Rapid object detection using a boosted cascade of simple features. *Computer Vision and Pattern Recognition Co*
6. Vertegaal, R., Weevers, I. & Sohn, C. (2002). GAZE-2: An attentive video conferencing system. *Poroceedings of CHI Extended Abstracts '02,* 736-737.

[11] http://www.skype.com/en/

[12] https://hangouts.google.com/

APPENDIX A. Taking Notes or Playing Games? Understanding Multitasking in Video Communication

Factors influencing perceptions of, and attitudes towards multitasking behavior are investigated in this study. Together with Jennifer Marlow and Daniel Avrahami, I conducted the experiments during my internship at Fuji Xerox Palo Alto Laboratories (FXPAL). The following paper describes the experiment and its findings, accepted for the conference of CSCW'16. This study forms the prior research for parts 1 and 2 of this thesis.

Taking Notes or Playing Games? Understanding Multitasking in Video Communication

Jennifer Marlow
FXPAL
marlow@fxpal.com

Eveline van Everdingen
Vrije Universiteit Amsterdam
e.a2.van.everdingen@student.vu.nl

Daniel Avrahami
FXPAL
daniel@fxpal.com

ABSTRACT
This paper presents a detailed examination of factors that affect perceptions of and attitudes towards multitasking in video conferencing. We first report findings from interviews with 15 professional users of videoconferencing. Our interviews revealed the roles and potential link of technology and activity. We then report results from a controlled online experiment with 397 participants based in the United States. Our results show that the technology used for multitasking has a significant effect on others' assumptions of what secondary activity the multitasker is likely engaged in, and that this assumed activity in turn affects evaluations of politeness and appropriateness. We also show that different layouts of the video conferencing UI can affect perception of engagement in the meeting and in turn ratings of polite or impolite behavior. We propose a conceptual model that captures our results and use the model to discuss implications for behavior and for the design of video communication tools.

Author Keywords
Multitasking; Video mediated communication; Experiment; Video conferencing

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction styles.

General Terms
Human Factors; Design; Measurement.

INTRODUCTION
In a connected digital world, in which notifications, communication, and immediate access to information all fight for our attention, the temptation to multitask during face-to-face and online meetings is great. Certainly, the modern workplace is full of distractions that can be difficult to manage. Multitasking often results from external events and interruptions, or it can be self-initiated [1, 10, 11, 13].

CSCW '16, February 27-March 02, 2016, San Francisco, CA, USA

ACM 978-1-4503-3592-8/16/02…$15.00
DOI: http://dx.doi.org/10.1145/2818048.2819975

Figure 1. Multitasking on a mobile device during a video conference (a keyframe from the experiment).

The pressure to multitask can be exacerbated by a workplace culture that requires employees to attend to multiple tasks or multiple projects at once [17]. With the proliferation of personal smartphones, tablets and other mobile devices in the workplace, opportunities to self-interrupt, be interrupted, or multitask have grown. Furthermore, with such personal devices, the delineation between personal and work activities in the workplace is blurring.

As distributed work teams become more prevalent, more and more meetings are being held over video conferencing. This can involve groups of people communicating via conference rooms with a specialized video setup, but another common scenario is one-on-one meetings with participants connecting to a video call from a personal laptop or desktop computer. This means that in some cases people are already using the computing device that they will also be using for multitasking.

While the effects of multitasking in face-to-face meetings have been previously studied, the perceptions and effects of multitasking in small, video-mediated work conversations – the focus of this paper – are not well known. Past research has provided anecdotal evidence that participants in video-mediated meetings and conversations both engage in multitasking behaviors and suspect others are doing so too (e.g. [6]). Building on these observations, the goal of the present work was to answer the following research question:

What factors influence perceptions of and attitudes towards multitasking behavior in video conferencing?

To investigate this question, we first interviewed people whose jobs involved remote collaboration about their experiences with multitasking during video conferences. These interviews uncovered parameters that influence the obviousness and acceptability of multitasking behavior. We then conducted a controlled experiment to tease apart the relative influence of various technical and social factors on perceptions of multitasking behaviors in a video conference-based meeting between two people (see Figure 1).

Through the combination of interviews and a large controlled experiment we make the following contributions to the field of CSCW: 1) An exploration of the context of multitasking in videoconferencing, uncovering the roles and potential link of technology and activity, 2) A controlled investigation of the effect of technology-related factors on observers' evaluations of multitasking behavior, and their assumptions about what secondary activity is being engaged in, and 3) A demonstration of the role that the design of the videoconferencing tool itself may have in influencing perceptions and outcomes. Our results have direct implications for the design of videoconferencing tools and device ecosystems for reducing the negative impact of multitasking on communication.

RELATED WORK

Multitasking in face-to-face meetings

Multitasking activities during a meeting can take many forms: they may be related to something that is going on in the meeting (e.g. verifying information, looking at the document being discussed), they may be related to other work that the participant needs to get done, or they may be personal and unrelated to either the meeting or work. Several lines of work have looked at the role of multitasking during in-person meetings.

One relevant area pertains to differences in the perceived politeness and self-reported frequency of using technology such as laptops and mobile phones for work or non-work purposes during meetings [5, 25]. Particularly, multitasking may not always be easy to detect. For example, when laptops are used during a meeting, speakers cannot be sure what a laptop is being used for (e.g. taking notes or browsing the internet), which may deter them from attributing disrespect [14]. Additionally, it is not always clear to an observer whether a multitasking behavior (for example, replying to a text message during a meeting) is necessary or not [17].

Finally, individual characteristics also influence the degree to which multitasking is viewed as acceptable. For example, an individual's own tendency to multitask [4], national culture [8], social and organizational norms [23] and gender [25] can influence perceptions of the acceptability and civility of doing other things during a meeting.

Video-mediated multitasking

Multitasking behavior over video conferencing has been described in both work-related [6] and personal conversations. For example, teenagers often divide their attention between multiple activities (like browsing the internet or social networking sites) while on a video call [7].

Several elements may contribute to a desire to multitask in online meetings: First, the communication itself most likely takes place on a computing device (a computer, tablet, phone, etc.) where other applications and information await. Next, unlike in face-to-face meetings, in video conferencing, only a small part of the user's body and surroundings are typically visible to other parties in the meeting allowing to potentially hide their secondary activity. Finally, it is even possible that participants would feel that in online meetings, even if it were known that they are multitasking it could be less obvious to other parties who are not co-located what it is they are doing.

One impact that multitasking can have on video communication is that it potentially affects gaze. People are fairly accurate at detecting when eye contact/gaze is being directed at them, and this feeling of direct eye contact can build trust [3]. For this reason, much research on building video conferencing systems has focused on finding ways to enhance the perception of direct eye contact (e.g. [12]). Multitasking can disrupt the feeling of direct eye gaze by causing an individual's attentional focus to be directed elsewhere. For example, when engaging in an unrelated chat or email during a video conference, attention and eye gaze will be shifted away from the webcam or video chat screen.

Therefore, the technology used for multitasking may play a big role in influencing perception. For example, a user performing the secondary task on their smartphone will appear different from a user performing the secondary task on the same screen where the videoconferencing software is running. People participating in meetings with a single device have greater obscurity with regards to what they are doing (they could be doing different things in different browser tabs, only one of which is visible at any given time.)

In addition to *where* a person is looking, the *reason* for the multitasking, or perhaps more importantly, the *reason assumed* by the other person, is likely to also influence perceptions of the appropriateness of the behavior. The effects of gaze are at least partly determined by a subjective interpretation [9]. Therefore, an observer's assumption about what a person is looking at could affect perceptions of behavior more than any absolute activity.

Consequences of multitasking

Multitasking in a meeting can be disruptive and counterproductive to other participants. For example, 51% of respondents to a recent survey about videoconferencing cited people who are multitasking – e.g., tapping on a keyboard – as a major distraction [21]. When there is a large group, there may be some expectation that not all participants need to be actively engaged. In one-on-one interdependent conversations, however, multitasking by one party all but guarantees a disruption in the flow of events [17].

The concept of actor-observer asymmetry [19] states that people often perceive their own behavior in a way that is different than how others perceive that behavior. This has been observed in the context of multitasking during face-to-face meetings [5], where respondents thought others were much more likely to be playing games, browsing their social network or browsing the web during meetings than they reported doing themselves.

Although individuals may think their own multitasking behavior is either unobtrusive or appropriate, they tend to evaluate others who are multitasking more harshly and see that behavior as rude [16]. Krishnan et al. [17] found that when a negotiation partner was receiving and checking text messages during a conversation, they were perceived as less professional and less trustworthy. It is possible that a similar phenomenon will also occur in video-mediated communication with regards to the perception of multitasking. A multitasker in a videoconference might also feel that their behavior is less obvious since their conversation partner does not have access to the full set of cues about what they are doing.

Ultimately, multitasking in a video conversation involves a series of tensions between a person's need or desire to perform multiple tasks, and a need to focus (and appear focused) on the video communication. The ambiguity of behavior that comes with multitasking in video conferencing could be beneficial or harmful to different parties in the conversation.

In this paper we examine the following general questions: *How is technology used for multitasking in video conferencing? What factors affect attitudes towards multitasking in video conferencing?*

To answer these questions, we first conducted a series of interviews with remote workers about their experiences with multitasking in video communication. We then used the insights gained through these interviews to design and conduct a controlled experiment that explores these factors in detail. In the remainder of this paper, we report the findings from the interview and experiment, then discuss implications for behavior and technology design.

VIDEOCONFERENCING IN PROFESSIONAL USE

To begin understanding the different instances in which multitasking might occur during video meetings, we conducted semi-structured interviews with professional users of videoconferencing technologies. These interviews were conducted as part of a larger study on videoconferencing practices. We interviewed 15 knowledge workers (11 male, 4 female) from a variety of industries and companies. In these interviews, participants discussed their use of video conferencing for both one-on-one and group video conferencing meetings. Across the cases, multitasking naturally emerged as a major theme when experiences with videoconferencing were discussed.

Participants mentioned engaging in multitasking themselves and suspecting that their meeting partners were doing the same. However, all multitasking was not the same, nor was the way in which it was performed. From the many examples given by the interviewees, the *nature* of the multitasking activity and the *device* on which the secondary activity occurred emerged as two key dimensions. We analyzed all examples pertaining to multitasking with these two dimensions in mind.

The nature of the multitasking activity

Comments about the multitasking activities fell into three main categories: First, the multitasking could be meeting related where a participant may wish to look at related documents, check facts, or engage in "sidebar conversations" related to the topic being discussed in the meeting. This type of behavior is related to the topic being discussed in the meeting, but may mean shifting one's attention to a different part of the screen, a different area of the screen than the video window, or to a different screen altogether. The second category of activities was other work-related multitasking in which a participant may be distracted by another important or urgent work-related inquiry, such as an IM message from another colleague or client. Finally, a third category was personal activities such as checking personal email, browsing Facebook, or playing with physical toys.

Meeting-related multitasking

Interviewees provided examples of cases where they (or their teammates) would multitask during meetings, but the multitasking behavior was relevant to the content being discussed. Typically this took place in large group conferences; this type of "sidebar" behavior has been frequently observed in other contexts as well (*c.f.*, [20]). As one interviewee said:

> *One thing that happens sometimes is we have IM chats, like side conversations going on during the meeting. Usually, I can't do that, because I am the chair, so I am usually sharing, so I can't.* (P9)

In cases where screen sharing of common referents like documents, photos, etc. was a main focus of activity, it was possible that a meeting participant would pull up the referent in a new browser tab or window and focus their attention on the artifact. This, however, was challenging because it meant that they were unable to focus on the video window showing the "talking head" of their conversation partner, but they were nevertheless engaged in the topic of the conversation even if their gaze was not directed at the other person.

> *I find it hard to engage [over videoconferencing] with people. I find I am not actually looking at them most of the time when I am video conferencing. It is the same way with the lesson. I am looking at the photo that he sent me, I am not looking at him.* (P7)

> *We weren't screen sharing or anything. I guess we just tabbed away from the video screen and then had the doc up. When we were referencing it, we could go through it together.* (P1)

Other work-related multitasking

Another category of multitasking behavior was dealing with disruptions related to other work not relevant to the meeting. For example, responding to another colleague's messages or interruptions during a meeting (for example, if a question needed to be addressed). This meant that the focus would not completely be on the conversation.

> *Sometimes I have other interaction going on with other team members and side comments. I have another IM going on.* (P2)

This type of behavior would sometimes occur in videoconferencing meetings that involve multiple people at a single location:

> *There's two people remotely, three people in the room, and somebody asks a question and the people remotely are talking, and then somebody in the room thinks it's okay to ask me a question about something slightly different while they're talking.* (P4)

Personal and unrelated multitasking

Finally, a third category of multitasking behavior that was mentioned was doing personal or unrelated multitasking during a meeting, often not paying attention to the conversation.

> *If you're not in on the conversation, if it's something completely unrelated to you, you might be playing with Buckyballs or reading your email instead or something like that.* (P11)

In fact, being able to multitask without being seen was cited as a reason for preferring to meet over the phone instead of via a video call:

> *Usually, I don't like doing video because if you're working on something else, it would be rude to not look at the camera all the time, so personally I prefer voice calls...* (P6)

The same interviewee noted that when she was working on something else off-screen, audio cues could still give away the fact that she was otherwise engaged:

> *If I'm typing something, you can hear it in the chat, and I feel really bad because it's obvious sometimes.* (P6)

Technology used for multitasking

Another theme that emerged was that participants had different technical configurations for their work. The comments that they made suggested that certain configurations were more likely to be used for some purposes than for others. For example, laptops could be used for meeting-related multitasking:

> *If something's mentioned mid-conversation, I might actually look into it while other people are talking about it on my laptop but yeah, it's not consistent amongst all the employees whether we do or not... [Name], for instance, almost always brings her laptop to the table and is usually multitasking in the middle of the conversation.* (P11)

Having multiple monitors also enabled one participant to refer to relevant meeting-related material in the course of a conversation:

> *I actually have four monitors on my system, and most people have two. You can be sharing a couple of screens and looking at a third screen to look up stuff.* (P9)

Other activities were not tied to an exclusive configuration. For example, taking notes or accessing email could occur on either a laptop *or* a mobile device:

> *Yes, usually take notes on my iPhone or on Note.* (P7)

> *I'd say 25% of the time I am doing emails on my laptop and I'd say 70% of the time or 65% of the time I'm doing it on my phone. Then 5 to 10% of the time I'm doing it on my iPad.* (P5)

These initial findings give valuable insight into self-reported multitasking behaviors and the range of activities engaged in, and devices or technologies used. These insights also expose three important questions: How does the technology used for multitasking affect observers' assumptions about the secondary activity? How does technology affect evaluations of the appropriateness and politeness of the multitasking behavior? And finally, does the video conferencing tool itself influence perceptions of multitasking?

EXPERIMENT

To further understand how the factors elicited in our interviews influence perceptions of, and attitudes towards multitasking behavior in video conferencing, we designed a controlled experiment. In the experiment, participants viewed one of a set of short video clips depicting a part of a videoconferencing meeting between two coworkers. In the clip (except for a control condition), one of the coworkers engaged in multitasking (see Figure 1).

Participants then rated the behavior of the people in the clip along several dimensions of interest, including politeness, engagement, and the likely secondary activity of the multitasker. We focused on one-on-one interactions as a good starting point because this is a setting in which perceived attention should be most important (there is no plausible deniability or doubt about participation). We were interested in what influenced the perception of *what activity a person was likely to be doing*, *how engaged did they seem*, and *how appropriate or polite their behavior was.* The two key elements manipulated in the experiment were thus the technology used for the secondary task, and the layout of the videoconferencing tool.

Hypotheses

In our interviews, laptops, second monitors, and smartphones were all mentioned as technologies used for multitasking during videoconferencing. Given that these devices are commonly associated with different primary uses, we hypothesize that seeing a person multitask on different devices will influence what observers assume the person is doing:

H_1. The technology used for multitasking will have a significant effect on the assumed secondary activity.

The assumed activity should, in turn, affect how polite and acceptable multitasking behavior is judged:

H_2. The assumed secondary activity will have a significant effect on ratings of politeness and acceptability.

Furthermore, the technology used may make multitasking easier to spot and may have a direct or indirect effect on how polite or acceptable the multitasking behavior is judged:

H_3. The technology used for multitasking will have a significant effect on ratings of politeness and acceptability.

Still, if H_3 is supported, would more overt multitasking be considered more or less acceptable? On one hand, multitasking on a different screen or device may be seen as more forthcoming and less deceiving, and thus more polite. On the other hand, multitasking on the same screen may be seen as less egregious and more ambiguous. We thus pose two, contradicting, sub-hypotheses for H_3, expecting at least one to be rejected.

H_{3a}: Multitasking on one screen will be rated as more polite and acceptable than on a second-screen or mobile device.

and

H_{3b}: Multitasking on a second-screen or mobile device will be rated as more polite and acceptable than on one screen.

Finally, we were interested in whether the presentation of the parties in the videoconference influences attitudes towards multitasking. While in some videoconferencing tools the local and remote participants are both shown at the same size, in others, the video of the remote participant occupies the majority of the screen and the local participant is only shown a small video of herself. We hypothesize that a larger view of multitasking by a remote participant would reduce ratings of their engagement in the meeting and the acceptability of their behavior:

H_4: Seeing multitasking in a visually larger format will reduce ratings of the multitasker's engagement in the meeting and how polite and acceptable their behavior is.

Method

For the experiment, which employed a between-subjects design, we created a set of short 1-minute video clips depicting a part of a meeting between two coworkers. In these clips, one coworker ("Person A") describes three different locations in which an advertising campaign could be conducted. The second coworker ("Person B") passively responds to Person A's remarks with short comments such as "yeah" and "uh-huh". While the clips varied along several dimensions (detailed below), all clips without exception followed the script shown in Figure 2.

The video clips were created as follows: First, in order to control for differences that might stem from the gender of the coworkers depicted in the videos, we created two versions of each clip: one in which Person A and B were both male, and one in which Person A and B were both female. The script, timing, and behaviors were identical across the two versions. For simplicity, in the remainder of this section, whenever a clip is described, it should be assumed that two identical clips were created for both actor genders.

In order to maintain experimental control, we recorded the video of Person A once, and used this video in all the conditions. This novel approach prevents any difference in the behavior of Person A from influencing participants' judgment of the behavior of Person B.

To test the effect of the technology used for multitasking, we created three versions of the video with Person B multitasking (see Figure 3) and a fourth version with no multitasking. In all conditions, Person B begins by directing their gaze towards the webcam (straight ahead). However, after a short notification sound is heard, Person B starts engaging in multitasking. In the first version, the *Same Screen* condition (Figure 3a), Person B directs their gaze to a different window on the same screen on which the videoconference is taking place. In the *Dual Monitor* condition (Figure 3b), Person B directs their gaze to a secondary monitor set up next to the main screen. In the *Mobile device* condition (Figure 3c), Person B directs their gaze downwards to a mobile phone they were holding and interacting with in their lap.

A:	...Which is why we need it by next Thursday.
B:	I see… Should I talk to Don about it?
A:	Yeah. That's a good idea.
A:	So...the main thing I have for today is that I wanted to get your opinion on three potential locations where we ***[notification sound plays]*** could show the ads as part of the marketing campaign
B:	Uh-huh
	B begins multitasking
A:	So, looking at the pictures I took today, the first location is a bus stop downtown near the courthouse. The Northbound and eastbound buses all stop here so there should be lots of potential traffic.
B:	Mmmhmm….
A:	The second location is a billboard that is next to the highway
B:	Yeah
A:	The third location is near the baseball stadium
B:	Sure
A:	Now, if you look at the email I sent, you can see the relative costs and potential views for each of the locations.
	B looks back at A
A:	The bus stop is going to be the least expensive, the billboard is the most expensive, and the baseball stadium is somewhere in between.

Figure 2. Video script: An exchange between two coworkers (Person A and Person B).

a) Multitasking on the same screen **b) Multitasking on a dual monitor** **c) Multitasking on a mobile device**

Figure 3. Sample key-frames showing B's gaze when multitasking in three experimental conditions.

a) Side-by-Side UI layout **b) Big-and-Small UI layout**

Figure 4. Two UI layouts used in the study based on common layouts in video conferencing applications.

In order to test hypothesis H_4, we created two versions of each video that varied in the presentation layout of the videoconferencing (see Figure 4). We chose two common layouts of videoconferencing applications. In the first layout, which we refer to as *"Big-and-Small"*, Person B is shown in a large window occupying the majority of the video and Person A appears in a small thumbnail (Figure 4b). This layout is similar to popular videoconferencing applications such as Skype and Google Hangouts. In the second, *"Side-by-Side"* layout, Person A and Person B are shown side by side with equal size video feeds (Figure 4a). This common videoconferencing layout is used most often in multi-person video meetings.

To summarize, the full experimental design was as follows:

× 4 Multitasking Technology
(*Same Screen, Dual Monitor, Mobile Device, None*)
× 2 Layout (*Side-by-Side* vs. *Big-and-Small*)
× 2 Actor gender (*Female-Female* vs. *Male-Male*)

= 16 conditions in total

Measures

After watching the video, participants were asked to give open-ended responses describing separately what they saw each person in the meeting do. We used these open-ended responses both to confirm that participants had paid attention to the video clip and also to get an initial sense of their impressions of and reactions to Person A and Person B's behaviors.

To test hypotheses H_2 and H_3, we asked participants to rate the behavior of Person B (the multitasker). Since the participants in our experiment only observed the meeting rather than participated in it (not an uncommon situation in multi-person meetings), we asked them to rate the obviousness of the multitasking from their perspective, and the perspective of Person A. The following were rated on a 5-point Likert-scale ranging from "Completely Disagree" to "Completely Agree"

1. Person B's behavior was *polite* during the conversation
2. Person B's behavior was *acceptable* during the conversation
3. It was obvious to me that Person B was multitasking
4. It was obvious to Person A that Person B was multitasking
5. Person A was engaged with the meeting
6. Person B was engaged with the meeting

Next, in order to test H_1 and H_2, participants were asked to rate the likelihood that Person B was engaged in each of a set of 8 secondary activities on a 5-point Likert-scale, from "Very Unlikely" to "Very Likely". Two of the 8 activities were meeting related (*"Focusing on the document person A is talking about"*, and *"Searching for other locations for the advertising campaign"*), three were related to other work activities (*"Looking at a document for another project"*, *"Reading a work-related email from another co-worker"*, and *"Chatting with another co-worker"*), and three were personal activities (*"Browsing Facebook"*, *"Reading a personal email"*, and *"Chatting with a friend"*). The order of the list of activities was randomized for each participant.

Finally, we collected demographic information, including participants' age, gender, experience with video conferencing (Skype, Google Hangouts, etc.) and comfort with multitasking (using the PAI scale of polychronic-monochronic tendency [15]).

Procedure

The experiment was set up as an online experiment using Amazon Mechanical Turk. Participation in the task was limited to individuals in the United States. Upon accepting the task, participants were randomly assigned to one of the sixteen conditions described above. Participants were shown a 1-minute video clip of the videoconferencing meeting based on their assigned condition. We used a timer on the page to prevent participants from advancing past the video before the video was finished. Participants then answered the open-ended questions, rated the behavior they saw in the video, and finally provided biographical data and information about their multitasking preferences.

We recorded the time participants took to interact with each page of questions in order to filter participants who quickly answered all the questions or neglected the task for very long durations. We also included a "check question" about the locations discussed in the meeting (see Figure 2) to ensure that participants had paid attention to the content of the video. The full task took 3 minutes to complete. Participants were paid 30 cents for their time.

Participants

436 participants completed the task[13]. We excluded 18 participants who failed to answer the check question correctly or gave the same rating value to all questions. We also excluded 10 participants whose completion time was too short/long based on a Mahalanobis Outlier analysis. Thus, our full dataset for analysis included 408 participants, with an average of 25 participants per condition (Min=17, Max=36, SD=5.36). 48% of participants were women. Age was reported in bands with 50% between the ages of 25-34.

RESULTS

Overall, participants appropriately rated Person A as highly engaged in the meeting (M=4.4) and Person B as less engaged (M=2.6). A Wilcoxon Signed Rank test shows this difference is significant (p<.001).

Politeness ratings and Acceptability ratings were significantly highly correlated (r=.82; p<.001). Ratings of Person B's *engagement* in the meeting were correlated with the behavior seen as polite (r=.71; p<.001) and acceptable (r=.67; p<.001). On the other hand, the more *obvious* it was that multitasking was taking place, the less polite (r=-.49; p<.001) and less acceptable (r=-.42; p<.001) the behavior seemed. A comparison showed that participants thought it was more obvious to them that multitasking was taking place than to Person A (M=3.5 vs. M=3.0; $t(396)=9.4$; p<.001). In the remaining analyses, we look only at how obvious multitasking was to the participant.

Open-ended responses

Checking participants' open-ended responses to confirm our manipulations, it was apparent that overall, they perceived what was happening in the videos as intended. Many expressed an opinion that Person A was "*actively involved in the conversation and was leading it.*" For the multitasking conditions, many responses referred to Person B doing something else (whether it was explicitly noted or 'presumed' behavior). Participants also referred to noticing a change in Person B's gaze, body language, and behavior (e.g. "*Person B started off engaged and then kind of got distracted or drifted off"*).

In the process of examining the open-ended responses, we also discovered that while participants in the Mobile Device condition easily identified that multitasking was taking place (and gave a high rating for multitasking being obvious, M=3.88), more than half of them thought that Person B's secondary activity was *taking notes* rather than interacting with their phone (as opposed to the Dual Monitor condition where multitasking was both rated as obvious (M=4.1) and correctly perceived as happening on a second screen). The fact that a downward gaze was as easily interpreted as being fully engaged in the meeting as it was as being completely disengaged is interesting. However, in order to understand attitudes towards multitasking in the context of technology use, we wanted to focus on exploring participants' reactions when they realized that Person B's activity was taking place on a phone, not on paper.

To address this, we recruited 49 new participants for the Mobile Device condition, this time clarifying that Person B's downwards gaze was directed at a phone. Specifically, prior to watching the video, instructions now stated that "During the meeting, one person is looking at their phone." Examining the responses of these new participants showed that while multitasking was rated just as obvious as before (M=3.87), ambiguity was reduced, with comments often referencing the use of a phone (e.g., "*Person B was distracted and checking his phone messaging. He wasn't offering any feedback to Person A except for a minimal 'mmhuh' etc. He was being rude, disrespectful and inattentive to Person A*", and "*She basically listened to Person A and then she started looking at her phone and not really listening. She was kind of rude*."). In subsequent analyses, for the Mobile Device condition, we used the data from these new participants. The full data analyzed thus includes a total of 397 participants.

Multitasking vs. No Multitasking

To test the underlying assumption that multitasking is perceived more negatively than no multitasking, we align-ranked participants' ratings [26] and performed analyses of variance (ANOVA) of Multitasking Technology on Politeness, Engagement (of B), and Obviousness of the

[13] Nearly half (47%) of participants attempted to complete the task a second and third time; however, to keep the experiment as a between-subject design, only their first attempt was kept.

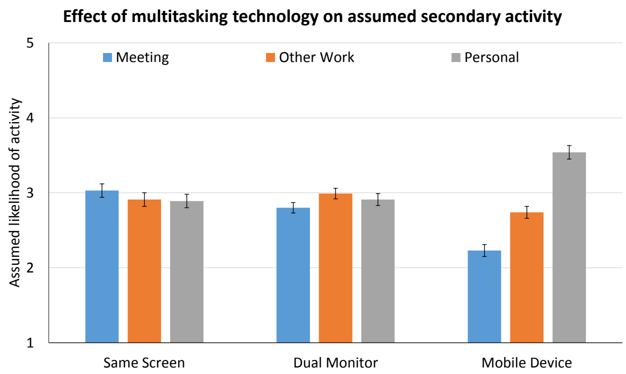

Figure 5. Effect of technology on assumed activity.

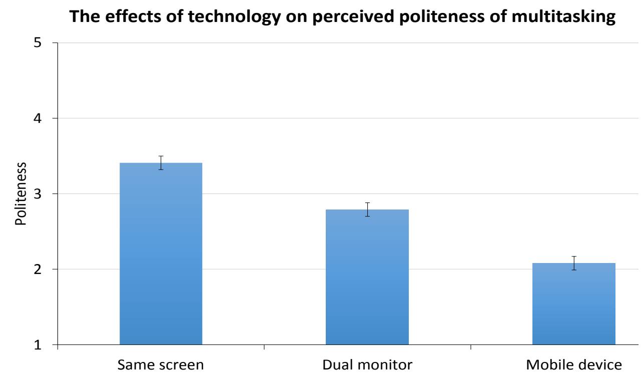

Figure 6. Effect of technology on politeness of multitasking.

multitasking. We then compared ratings from the Control condition and each other condition.

As expected, participants in the Control condition rated the Obviousness of multitasking very low, significantly lower than any of the multitasking conditions (M=2.3 vs. M=3.9; F[3,393]=50.7; *p*<.001). Behavior in the Control group was also rated as more polite (M=3.8) than in any of the other conditions (F[3,393]=53.5; *p*<.001) and more acceptable (M=3.7) than the Dual Monitor and Mobile Device conditions (F[3,393]=36.4, p<.001). However, engagement in the meeting in the Control condition was rated significantly higher than the Dual Monitor and Mobile Device conditions, but not different from the Same Screen condition (F[3,393]=18.7; *p*<.001).

We were now ready to test our hypotheses. In the analyses described below, we describe results from the 285 participants in the multitasking conditions.

The effect of technology on assumed secondary activity

To test H_1, which stated that the technology used for multitasking will affect the assumed secondary activity, we first combined the list of eight secondary activities into three Activity categories: *Meeting-related* activities, *Other-Work-related* activities, and *Personal* activities. We used the average likelihood rating provided by each participant for each category. We then conducted a repeated measures mixed-model ANOVA on the align-ranked ratings, with Activity Likelihood as the dependent measure. We included Activity Category, Multitasking Technology, and the interaction between them as independent measures. The gender of people in the video ("Video-Gender"), UI Layout, and the Participant Gender were included as control. ParticipantID was modeled as a random effect.

As shown in Figure 5, the interaction between Multitasking Technology and Activity Category was significant (F[4,573]=20.6; *p*<.001). When a mobile device was used for multitasking, participant assumed that the secondary activity was significantly more likely to be personal (M=3.54) than related to the meeting (M=2.23; *t*(573)=7.4; *p*<.001) or to other work (M=2.74; *t*(570)=2.4; *p*<.05). This result is important given that, according to the findings of our interviews, phones are often used for meeting related tasks. In the Same Screen condition, on the other hand, meeting-related activities were assumed to be most likely (M=3.03), significantly more than personal activities (M=2.89; *t*(574)=4.3; *p*<.001). Specifically, it suggests that people will assume that someone multitasking on their phone is doing something personal and unrelated to the meeting, while someone multitasking on the same screen is likely doing something related to the meeting. **This confirms H_1.**

Looking at all eight activities for more detail, we again find a significant interaction between the technology used and how likely a secondary activity was rated (F[14,2017]=16.9; *p*<.001). The activity assumed most likely for the Mobile Device condition was "Chatting with a friend" (M=3.8). For the Dual Monitor condition it was "Reading a work-related email from another co-worker". "Focusing on the document person A is talking about" was the most likely assumed activity in the Same Screen condition, significantly higher than in either the Dual Monitor or Mobile Device conditions.

Effects on attitudes towards multitasking

To test our hypotheses on factors influencing attitudes towards behavior, we performed analyses of variance (ANOVA) with Politeness, Obviousness, and Engagement as dependent measures and Multitasking Technology, UI Layout, Video-Gender, and Participant Gender as independent measures. To test the effect of assumed activity, we also included participants' likelihood ratings of activities (Meeting-related, Other-work-related, and Personal). Having found no effects of Age-range, experience with videoconferencing, or Polychronicity tendencies, those were not included in the models reported here.

Effects on Politeness and Acceptability

Testing our second hypothesis, our analysis found a significant effect of the assumed secondary activity on Politeness. The more participants assumed the secondary activity was Meeting-related, the more polite they rated the behavior (F[1,276]=40.9; *p*<.001) and the more acceptable (F[1,276]=57.6, *p*<.001). The more they assumed the activity was Personal or related to Other Work, the less polite they viewed the multitasking (F[1,276]=30.0; *p*<.001 and F[1,276]=7.8; *p*<.01). The more the secondary activity was assumed to be Personal, the less acceptable it was rated (F[1,276]=40.72, *p*<.001).

As shown in Figure 6, the technology used for multitasking also had a significant main effect on Politeness (F[2,276]=18.5; *p*<.001). Technology used also had a main effect on Acceptability (F[2,276]=14.7, *p*<.001). A post-hoc comparison showed that Same Screen was rated as significantly more *polite* (M=3.4) than both the Mobile Device (M=2.1; *t*(276)=5.8, *p*<.001) and Dual Monitor (M=2.8; *t*(276)=3.97, *p*<.001). We found no additional effects on Politeness. Similarly, the Same Screen was also rated as more *acceptable* (M=3.5) than both the Mobile Device (M=2.19; *t*(276)=24.04, *p*<.001) and the Dual Monitor (M=2.88, *t*(276)=17.67, *p*<.001).

These results confirm both H_2 and H_3 and suggest that the technology a person uses for multitasking, and what people assume they are doing with that technology, significantly affect how polite and acceptable that behavior is judged.

Effects on Obviousness of multitasking

Looking at the effect of the technology used for multitasking on how Obvious the multitasking was, we find a significant effect (F[2,276]=3.7; *p*<.05), with multitasking rated as significantly more obvious on the Dual Monitor than on the Same Screen (4.1 vs. 3.6; *t*(276)=2.64, *p*<.001). Since less obvious multitasking (Same Screen) was also associated with higher politeness ratings, we may **accept H_{3a} and reject H_{3b}.**

Effects on Perceived Engagement

Hypothesis H_4 predicted that seeing Person B multitasking in a visually larger format will make them appear less engaged in the meeting and, in turn, less polite. Our analysis found a main effect of UI Layout on ratings of Engagement. However, the effect was the opposite of what we expected; Person B was rated as significantly less engaged when viewed in the *Side-by-Side* layout (M=2.3) than in the *Big-and-Small* layout (M=2.8; F[1,276]=15.8; *p*<.001). Thus, **we must reject H_4**. We return to this surprising result in the Discussion section.

Perceived engagement was also affected by the technology used for multitasking (F[2,276]=4.3; *p*<.05). Engagement was rated significantly lower when multitasking on a Mobile Device compared to on the Same Screen (2.0 vs. 2.9; *t*(276)=2.9; *p*<.01).

Effects of personal characteristics

As mentioned earlier, we found no effects of Age-range, experience with videoconferencing, or Polychronicity tendencies. We also analyzed the effects of Video-Gender as well as the participant's gender on attitudes and perceptions.

We found no significant differences between participants that watched a meeting between the two women and a meeting between the two men. We did find a small significant effect of the participant's gender on ratings of obviousness, with women perceiving multitasking as more obvious than men (4.0 vs. 3.7; F[1,276]=4.5; *p*<.05).

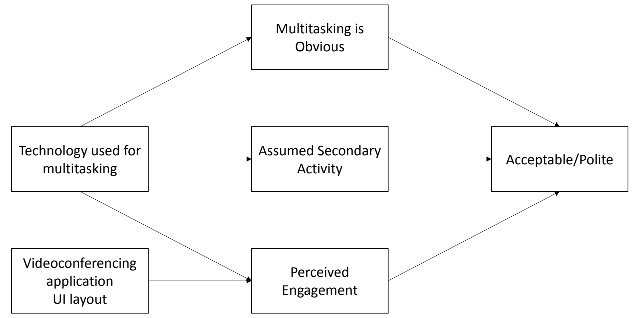

Figure 7. A proposed model of factors influencing perception of, and attitudes towards multitasking in video conferencing.

DISCUSSION

Our interviews and experiment give insight into multitasking during videoconferencing from the perspective of both the multitasker and the observer of multitasking. Our study also provides a first examination into how the unique affordances of video conferencing (such as UI layouts which show participants in different sizes, or a field of view that restricts what is seen) influence how such behavior is seen. In Figure 7, we propose a model summarizing the role of technology and UI layout on our outcome measures.

The results reveal that observers of video-mediated meetings use cues such as the technology used for multitasking to form assumptions about the secondary activity that is taking place. This relates to findings from studies of the use of laptops in face-to-face meetings where multitasking was ambiguously interpreted (e.g. [14]). However, as demonstrated by our work, ambiguity in the case of videoconferencing can be significantly greater. As Table 1 illustrates, first, in video communication, the device used for multitasking may be outside the camera's field of view, as in the case of a mobile device. Next, when activities all take place on a single screen, there will often be ambiguity about whether multitasking is at all taking place. Finally, even if the location of multitasking is apparent, as in the case of a dual monitor, the secondary activity is still unknown (similar to the face-to-face case in [14,17]).

Secondary activity on:	Is multitasking obvious?	Is the technology used obvious?
Single screen	No	Yes
Dual Monitor	Yes	Yes
Mobile device	Yes	No

Table 1. Variations in ambiguity of behavior and technology used for multitasking.

As discussed in [6], participants in videoconferencing may hope that this ambiguity allows them to hide their multitasking, while at the same time believe they are able to detect multitasking in others. Indeed, our study participants believed it was easier for them to spot multitasking than for the person in the video.

Our findings also highlight how observers' assumptions about the secondary activity will affect their reactions to that behavior. For example, when multitasking took place on a phone, activities were assumed to be of a personal nature. This finding is important since the secondary activity is rarely entirely known to the videoconferencing partner, and an incorrect interpretation will lead to incorrect evaluation. Consider, for example, our interview findings where individuals reported using mobile devices also for meeting-related activities; in those cases, an entirely appropriate behavior may be judged negatively.

Additionally, the technology used influences how apparent, or obvious, the multitasking behavior is. Engagement was positively correlated with politeness/acceptability, while obviousness had the reverse effect. As we show, these assumptions and interpretations of the secondary activity can negatively affect how behavior is viewed as more or less polite, engaged, and acceptable. This has implications for the behavior of people who participate in video meetings, as well as the design of future video conferencing technologies.

Our experiment also revealed a surprising effect of the videoconferencing tool itself on how engaged meeting participants appear. Our comparison of two popular videoconferencing UI layouts showed that, contrary to our expectations, the multitasker appeared *more* engaged when shown in a large window, and *less* engaged when appearing in a smaller side-by-side view at equal size as the other meeting participant. There are at least two explanations for this effect: One explanation for this finding is that being able to see a participant in a larger window gives extra emphasis when their gaze is directed towards the camera. An alternate explanation is that when presented side-by-side, direct comparison between meeting participants can be easily made, with a multitasker appearing much less engaged than their counterpart. This result may have direct implication for the design of videoconferencing applications.

Implications for behavior

Our findings hint at how behavioral cues may highlight or mask multitasking behavior and perceptions thereof. Clearly, the best course of action is *not to engage in activities unrelated to the meeting at all*. However, if multitasking is unavoidable, our findings suggest that multitasking on the same device appears less inappropriate than interactions on a second screen or device. Multitasking on a dual monitor was easy to spot and rated negatively (although at least secondary activities were assumed to be work-related rather than personal).

The case of interacting with a mobile device during a videoconference is interesting because the device itself is not seen in the video in most camera configurations. Thus, the multitasker may benefit from their behavior misinterpreted as paying attention (e.g., taking notes). However, if activity is correctly interpreted as interaction on a phone, their interlocutors are likely to assume that they are engaged in an unrelated personal activity, such as playing a game, texting, chatting, etc. Two potential recommendations emerge: First, hoping to rely on the phone being off camera to hide multitasking may prove risky. Second, users may wish to tell their meeting partners when they are using their phone for meeting-related activities, to prevent those from incorrectly assuming the activities are personal.

Implications for design

Our findings offer several implications for the design of video conferencing systems. Eye gaze and body language being directed elsewhere besides (relatively) straight ahead at the camera, are easy to detect and negatively perceived by meeting viewers. Therefore, design decisions may vary depending on whether they are geared towards helping a multitasker be perceived less negatively, or helping another participant have a better idea of what their conversation partner is doing.

In the first case, systems should be designed such that users can direct all incoming alerts and notifications from their smartphone and secondary monitors to a primary screen of their choice while they are in "in a meeting" status. This way, disruptive gaze re-direction can be minimized.

Second, in the case of multiple monitors and multiple cameras, a system could be designed to use face-detection to dynamically determine which screen or camera a person is facing and automatically switch to that camera. This may allow users to maintain a consistent impression of direct eye gaze, translating the apparent benefit of the single-screen condition to the case of multiple monitors.

Finally, video conferencing systems could allow for the layout and information about participants' behavior to be displayed differently based on individual tendencies/characteristics. The video window could also be supplemented with additional cues about the other's behavior, but such an option should be treated with care. Visualizations of others' behavior or psychological states over video conferencing have been explored in other contexts such as visualizing browser activity [18], stress or mood levels [22], 24]. While these can be useful, there are well-documented social tradeoffs between increasing awareness of another person's activity and preserving privacy. Thus, more work is needed to know how cues related to meeting relevant and irrelevant behavior are perceived by both sides in a multitasking context where participants may have different motives and privacy concerns.

Limitations and future work

One limitation of our experiment is that participants were placed in the role of non-participating meeting observer watching two strangers converse. It is possible that participants' reactions or perceptions would be different if they were actively participating in a video meeting with people they knew well. It is important to note, however, that being a passive observer in a meeting is not uncommon, particularly in multi-person videoconferencing meetings. As noted by our interviewees, in multi-person video meetings, meeting participants will try to manage how their activities appear to others, and judge other participants' multitasking behavior. Furthermore, our results strongly suggest that participants understood the situation well; their open-ended responses often contained affective attributions stemming from the multitasking behavior.

Next, while our experiment focused on a handful of factors that emerged from our interviews (e.g., the technology used and the secondary activity), as we learned from our interviews, video meetings can take a variety of forms and be held between many different types of people. As such, other dimensions of multitasking in video meetings still need to be explored in future work.

One such dimension involves the nature of the meeting and the participants. As the number of meeting participants grows or the length of the meeting increases, attitudes towards one or several people's multitasking behavior could change. Social dynamics may also play a role: In our experiment, conversation dynamics suggested that one person was in charge and leading the meeting. Exploring additional status relationships could be important since this relationship may influence attitudes towards the multitasker (e.g. [25]).

Another interesting dimension to investigate is whether if a person verbally elevates their multitasking behavior to the foreground (e.g., by stating "*Hold on, I need to reply to this message.*"), their behavior could appear more appropriate or polite. Similarly, new technologies such as smartwatches and head-mounted displays are becoming available and present interesting new context for multitasking. Exploring the role of these and other variables in future work will give greater insight into additional factors (both social and technical) that influence the dynamics of multitasking in meetings held over video conferencing.

Finally, the videoconferencing itself, not only the secondary activity, can take place on a range of devices and setups, ranging from dedicated videoconferencing rooms to running on a mobile device. In particular when videoconferencing on a phone or a tablet, the ability and method for multitasking is different: applications typically occupy the entire screen, and since the device is completely mobile the camera also moves around, potentially breaking any illusion of eye contact. In future work we plan to explore the effect of multitasking in additional videoconferencing setups, and build and evaluate the effectiveness of tools that dynamically transition between cameras to follow the multitasker's gaze.

CONCLUSION

This paper addresses the tension between the growing use of video conferencing and meeting participants' desire for multitasking. Through a combination of interviews and controlled experimentation, we provide a detailed account of technological and video conferencing UI-based factors influencing attitudes towards multitasking in videoconference meetings. Given the nature of the modern workplace and the increasing proliferation of mobile devices, it is unlikely that workers' tendency for multitasking will go away soon. Our results indicate that all multitasking behavior is not perceived equally; the association of technology and common activities along with how the activity is presented through the video conferencing tool layout influences observers' assumptions and attitudes. While we do not pass judgment on whether or not multitasking should occur, our study provides insights into how the negative interpersonal perceptions that come along with it when it does happen can be accentuated or reduced using technology and social cues.

REFERENCES

1. Rachel F. Adler and Raquel Benbunan-Fich. 2013. Self-interruptions in discretionary multitasking. *Computers in Human Behavior* 29, 4, 1441–1449.
2. Robert Bajko. 2011. Mobile Telephone Usage and Perception During Group Meetings. *Proc. CONISAR,* 1–10.
3. Ernst Bekkering and J.P. Shim. 2006. Trust in Videoconferencing. *Commun. ACM* 49, 7, 103–107.
4. Caroline S. Bell, Deborah R. Compeau, and Fernando Olivera. 2005. Understanding the social implications of technological multitasking: A conceptual model. *Proc. SIGHCI 2005*, 2.
5. Matthias Böhmer, T. Scott Saponas, and Jaime Teevan. 2013. Smartphone use does not have to be rude: making phones a collaborative presence in meetings. *Proc. MobileHCI*, ACM, 342–351.
6. Jed R. Brubaker, Gina Venolia, and John C. Tang. 2012. Focusing on shared experiences: moving beyond the camera in video communication. *Proc. DIS*, ACM, 96–105.
7. Tatiana Buhler, Carman Neustaedter, and Serena Hillman. 2013. How and why teenagers use video chat. *Proc. CSCW*, ACM, 759–768.
8. Peter W. Cardon and Ying Dai. 2014. Mobile Phone Use in Meetings among Chinese Professionals: Perspectives on Multicommunication and Civility. *Global Advances in Business Communication* 3, 1, 2.
9. Elena Francesca Corriero, Stephanie Tom Tong, and Pradeep Sopory. 2015. Behaviors, Perceptions, Responsiveness, and Presence: The Dyadic Model of Mediated Communication. *HICCS,* 462-471.
10. Mary Czerwinski, Eric Horvitz, and Susan Wilhite. 2004. A diary study of task switching and interruptions. *Proc. CHI*, ACM, 175–182.
11. Laura Dabbish, Gloria Mark, and Víctor M. González. 2011. Why do I keep interrupting myself?: environment, habit and self-interruption. *Proc. CHI*, ACM, 3127–3130.
12. Jim Gemmell, Kentaro Toyama, C. Lawrence Zitnick, Thomas Kang, and Steven Seitz. 2000. Gaze awareness

for video-conferencing: A software approach. *IEEE Multimedia* 7, 4, 26–35.

13. Victor M. González and Gloria Mark. 2004. Constant, constant, multi-tasking craziness: managing multiple working spheres. *Proc. CHI*, ACM, 113–120.
14. Shamsi T. Iqbal, Jonathan Grudin, and Eric Horvitz. 2011. Peripheral computing during presentations: perspectives on costs and preferences. *Proc. CHI*, ACM, 891–894.
15. Carol Felker Kaufman, Paul M. Lane, and Jay D. Lindquist. 1991. Exploring more than 24 hours a day: A preliminary investigation of polychronic time use. *Journal of Consumer Research*, 392–401.
16. Lisa Kleinman. 2010. Physically present, mentally absent? Technology multitasking in organizational meetings. PhD dissertation, University of Texas.
17. Aparna Krishnan, Terri R. Kurtzberg, and Charles E. Naquin. 2014. The Curse of the Smartphone: Electronic Multitasking in Negotiations. *Negotiation Journal* 30, 2, 191–208.
18. Danielle Lottridge, Eli Marschner, Ellen Wang, Maria Romanovsky, and Clifford Nass. 2012. Browser Design Impacts Multitasking. *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, SAGE Publications, 1957–1961.
19. Bertram F. Malle. 2006. The actor-observer asymmetry in attribution: a (surprising) meta-analysis. *Psychological Bulletin* 132, 6, 895-919.
20. Bonnie A. Nardi, Steve Whittaker, and Erin Bradner. 2000. Interaction and outeraction: instant messaging in action. *Proc. CSCW*, ACM, 79–88.
21. Polycom report. Video Conferencing Expected to be Preferred Business Communications Tool in 2016 According to New Survey on Global Video Conferencing Trends and Etiquette (2013).http://www.polycom.com/company/news/press-releases/2013/20131015.html [Accessed May 2015].
22. Andreas Sonderegger, Denis Lalanne, Luisa Bergholz, Fabien Ringeval, and Juergen Sauer. 2013. Computer-supported work in partially distributed and co-located teams: the influence of mood feedback. In *Proc. INTERACT*. Springer, 445–460.
23. Keri K. Stephens and Jennifer Davis. 2009. The social influences on electronic multitasking in organizational meetings. *Management Communication Quarterly*.
24. Chiew Seng Sean Tan, Kris Luyten, Jan Van Den Bergh, Johannes Schöning, and Karin Coninx. 2014. The role of physiological cues during remote collaboration. *Presence: Teleoperators and Virtual Environments* 23, 1, 90–107.
25. Melvin C. Washington, Ephraim A. Okoro, and Peter W. Cardon. 2013. Perceptions of civility for mobile phone use in formal and informal meetings. *Business Communication Quarterly,* 52-64.
26. Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 143–146

APPENDIX B. Experiment Materials

MOVIES

The movies which are used in the experiment 1 and 2 of part 1 in the thesis can be watched online.

Experiment 1

Static View: https://www.youtube.com/watch?v=_0N4HeiWxcc

Dynamic View: https://www.youtube.com/watch?v=T717Kgp8Hw8

Control group: https://www.youtube.com/watch?v=7uRZtilBDVM

Experiment 2

Static View: https://www.youtube.com/watch?v=Jb573r0mpgU

Dynamic View: https://www.youtube.com/watch?v=WUO7bpaiEgA

Control Group: https://www.youtube.com/watch?v=LeOiUeBWI5w

DATA ANALYSIS

The raw data files, filtered dataset and participant distribution files are included in [http://dx.doi.org/10.6084/m9.figshare.1619896].

TECHNICAL CODE

The code and short clip of the performance of the gaze-following technology with two webcams can be accessed at [http://dx.doi.org/10.6084/m9.figshare.1619896].

ACKNOWLEDGEMENTS
I would like to thank my supervisor Victor de Boer for the support and insightful comments of my Master Project study. Also I thank Jennifer Marlow and Daniel Avrahami, who provided me an opportunity to join their team as an intern. I learned a lot from their expertise, without them I would not have come as far and with this level of quality of the research as I am now. I would like to thank my family: my parents and my sisters for their mental support and practical feedback on the invented two-webcam gaze-following tool. At last but not least, Annemarij de Groot deserves to be mentioned for her reviewing skills.