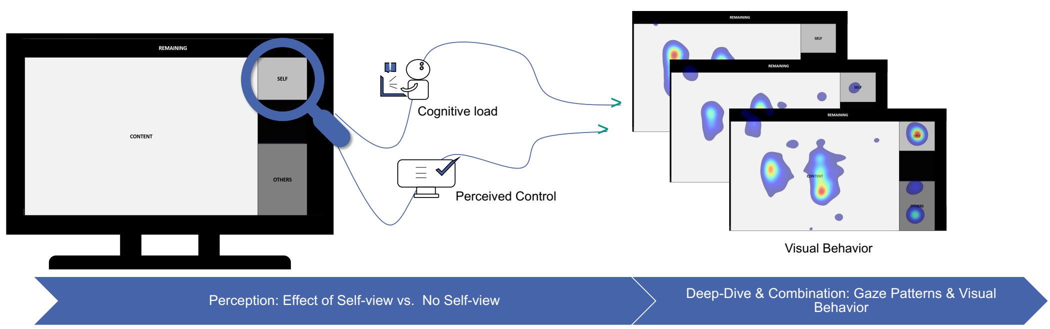
"Mirror, mirror in the call": Exploring the Ambivalent Nature of the Self-view in Video Meeting Systems with Self-Reported & Eye-Tracking Data

JULIA SEITZ, Karlsruhe Institute of Technology (KIT), human-centered systems lab (h-lab), Germany
IVO BENKE, Karlsruhe Institute of Technology (KIT), human-centered systems lab (h-lab), Germany
ALEXANDER MAEDCHE, Karlsruhe Institute of Technology (KIT), human-centered systems lab (h-lab), Germany



Video meeting systems offer great potential for work and life, but they can also have negative effects. One reason is the presence of technical stimuli that do not exist in the physical world. A prominent example is the self-view feature, a mirrored image of oneself shown during the video meeting. The self-view feature comes with a trade-off between the advantage of enhancing control and the disadvantage of increasing cognitive load of its users. So far, research is scarce when it comes to understanding this ambivalent nature and studies mostly relied on self-reported data without considering the actual interaction with the self-view. To address this gap, we conducted an experimental study with 57 participants and two design variants (with/without self-view), analyzed user perceptions through surveys and interviews, and explored gaze patterns using eye-tracking technology. Results reveal varying perceptions of cognitive load and control among self-view users and between the design variants, highlighting its ambivalent nature. We see differences in how participants interact with the self-view. In a cluster analysis, we identify three user groups (Benefiting Users, Cognitively Challenged Users, Control Losing Users). These groups also show differences in visual behavior, especially median fixation duration, and user characteristics. Based on our findings, we outline design recommendations for more flexible and intelligent design solutions by considering user groups and their individual differences.

CCS Concepts: • **Human-centered computing** → **Laboratory experiments**; **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**; **Social networking sites**; **Computer supported cooperative work**.

Additional Key Words and Phrases: self-view, video meeting systems, eye-tracking

Authors' Contact Information: Julia Seitz, julia.seitz@kit.edu, Karlsruhe Institute of Technology (KIT), human-centered systems lab (h-lab), Germany; Ivo Benke, Karlsruhe Institute of Technology (KIT), human-centered systems lab (h-lab), Germany; Alexander Maedche, Karlsruhe Institute of Technology (KIT), human-centered systems lab (h-lab), Germany.

ACM 2573-0142/2024/11-ART392
https://doi.org/10.1145/3686931

ACM Reference Format:
Julia Seitz, Ivo Benke, and Alexander Maedche. 2024. "Mirror, mirror in the call": Exploring the Ambivalent Nature of the Self-view in Video Meeting Systems with Self-Reported & Eye-Tracking Data. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 392 (November 2024), 37 pages. https://doi.org/10.1145/3686931

1 Introduction

Through the rise of remote work, video meeting systems (VMS) have become a key technology for organizations [16]. While VMS provide a lot of advantages for collaborating independent of the location, their intensive usage also comes with downsides for users. As an example, in comparison to in-person meetings, users experience a lack of social presence, lack of eye contact, video meeting fatigue, or unbalanced communication and understanding [4, 32]. Ultimately this may decrease the willingness to collaborate and the work performance [10, 50].

One potential driver to negative impacts of video meetings is the self-view feature [81]. This feature is a fundamental difference to in-person meetings in the physical world since the possibility to see yourself is not existent in traditional in-person communication. The self-view feature is a technical stimulus offered by VMS to provide an additional source of information in video meetings compared to in-person meetings. The self-view is automatically visible as a static depiction of oneself in a video meeting. Typically, VMS offer the opportunity to switch off the self-view [84]. Interestingly, the reasons why the self-view is present are not as clear as expected. Following the design principle of feedback outlined by Nielsen [70], one potential reason is providing the user with a control window on the information they sent to others [20]. Existing studies confirm this controlling nature of the self-view which seems to comfort the user [5].

On the other hand, previous studies have shown that the self-view requires allocation of cognitive resources. Thereby, it may lead to cognitive load, distraction, and over a longer period to video meeting fatigue [5, 28, 42, 69]. Complementing these findings, the results about the self-view's effect are diverging and seem to be ambivalent. It remains unclear how users are actually interacting with the self-view in case it is present [39, 42, 65].

In addition, next to the mere existence of the self-view stimulus, also other factors can influence how it impacts its users. So far, studies investigating the self-view explored its impact mostly in dyadic team settings and tasks which did not require using a shared screen setup, showing meeting content as well as the video feeds from other participants and the feed from oneself (see e.g., [39, 42]). In such a setting, the size of the self-view stimulus is very prominent. However, looking at daily meeting behavior, we see that a lot of meetings take place in small groups of three to five participants or even larger groups and in a shared screen setting (i.e., in webinars, educational settings, and similar). In such a meeting context, the self-view is depicted in a drastically smaller size compared to its depiction in a setup with two participants and no shared screen.

Our goal is to develop an in-depth understanding of the self-view and how it impacts its users in a video meeting with teams of more than two participants using a shared screen. To do so, we investigate the ambivalent nature of the self-view in VMS and focus on the perceived cognitive load and perceived control and aim to first answer the following research question:

(1) *"What is the effect of the self-view feature in VMS on users' perceived cognitive load and perceived control?"*

Analyzing the self-view feature and its effects in detail is challenging with subjective and perceptive measures only [52]. Taking a deeper look at the eye gaze provides insights into the actual visual behavior related to the self-view feature and can inform us about how users really interact with the self-view. Based on the eye-mind hypothesis, data collected via eye-tracking technology provides information about eye gaze, user's visual attention and underlying cognitive processes

[74]. This is especially interesting since faces seem to impact our visual attention differently than other objects by grabbing attention faster and more easily. Furthermore, the face of oneself is recognized faster than the face of others [48]. To the best of our knowledge, only few studies have investigated the gaze on the self-view and none by using eye-tracking technology. Therefore, we ask the following second research question:

(2) *"What is the visual behavior of users with regards to the self-view feature in VMS? Do different perceptions of the self-view impact visual behavior?"*

To answer these research questions (RQs), we present an eye-tracking between-subjects experimental study in the laboratory with 57 participants. We analyze the effect of established configurations of the self-view feature (self-view, no self-view) on the users' perceived cognitive load and control. In the study, participants performed three decision-making tasks in a sequence of three video meetings. We analyze user perceptions with self-reported quantitative measures as well as qualitative interviews, and user's visual behavior through eye-tracking. Based on these results, we create clusters of individual interactions with the self-view and differences in perception.

With our study, we contribute to the VMS body-of-knowledge by investigating both aspects of the ambivalent nature of existing configurations of the self-view in a single study and are the first to investigate visual behavior on the self-view holistically. We show differences in perceptions over a row of meetings. Most importantly, our study provides an in-depth examination of users' visual behavior when it comes to interacting with the self-view by highlighting differences in gaze patterns. We also combine the perception of the self-view with the different visual behaviors and identify groups of users that are more impacted and less impacted by the self-view. Based on these insights, we then outline recommendations for future self-view designs.

2 Conceptual Foundations

2.1 Video Meeting Systems and the Self-view

VMS, also called video conferencing systems, web conferencing or teleconferencing, have been studied for quite a while in human-computer interaction (HCI) and communications literature and early literature dates back to the 1980s [25, 54, 81]. Existing work on VMS does investigate the impact of the system on the user and on their productivity. Here, especially qualitative studies outline a broad range of problems and benefits of using the systems (e.g., [6, 40]). Due to the increased use of VMS in the last years, multiple new insights have emerged that are especially targeting well-being-related user states such as the construct of video meeting fatigue [22, 45].

VMS in general can be conceptualized as a communication technology that allows users to connect in real-time across different places [47]. Based on media naturalness theory and synchronicity theory, video-based communication is richer than text-based and audio-based communication [9, 55, 94]. However, still, it is not fully capable of mimicking a face-to-face conversation due to lacking cues [9, 55, 94]. By seeing each other, VMS allow users to better understand their partner's willingness and availability to communicate and to share additional information via screen sharing. Depending on the system used, additional features are included in the system and allow users to send files, chat messages, or make use of reactions (e.g., emojis or raising hands [2, 56, 94]). Looking at recent developments, especially since the pandemic, a wide range of features has been added to those contemporary VMS (see e.g., chatbot agents, or analytical dashboards as visible in the Zoom marketplace [96]). Besides, one feature that has existed in VMS from the start and discussed ever since is the self-view, a mirrored image of oneself.

Building on findings from early research focusing on the impact of seeing oneself in face-to-face interactions when using mirrors, the impact of the self-view was consequently also discussed in VMS [12]. To date, numerous studies have explored the impact on selected user states and meeting outcomes such as cognitive load, performance, or anxieties, like mirroring anxiety [1, 5, 19, 39, 42, 57, 65, 66, 93]. Thereby, the findings differ based on the intensity and the direction of the effect. Furthermore, individual characteristics such as the disposition of social anxiety or also long-term effects over time have been explored [19, 39, 42, 57, 65, 93]. More recently, its relation to the emerging phenomenon of video meeting fatigue has been highly discussed [1]. Thereby, especially based on qualitative findings, there seems to be an ambivalent mechanism of the self-view, subsequently called the ambivalent nature, leading to cognitive load and control.

2.2 The Ambivalent Nature of the Self-View: Its Effect on Cognitive Load and Control

The self-view seems to have an ambivalent nature, as depicted in Figure 1. According to our knowledge, existing literature so far always only considers one aspect of the ambivalent nature of the self-view. Below we explain both natures and summarize existing findings as well as the underlying psychological theories.

Nature 1: Self-view and Cognitive Load. As visible by Balogová and Brumby [5], the self-view seems to induce a distraction to users. This can be explained by the fact that additional stimuli lead to effort that needs to be spent on processing them, also called cognitive load. Cognitive load refers to the capacity of working memory used for performing a certain task [85]. Thereby, cognitive load theory is a very prominent theory explaining the concept. Originating from learning theory, it argues that cognitive load consists of effort due to a specific topic (intrinsic load), effort due to information presentation (extraneous), and effort due to creating a store of knowledge (germane) [85]. Based on the theory, additional elements displayed lead to higher extraneous load and may even limit the resources that can be spent on the task because of attentional conflicts in the working memory due to its limited capacity [13, 42]. The self-view thereby represents such an additional element when comparing video meetings to physical meetings but also to video meetings showing no self-view [78]. This leads to the conclusion that a constantly available self-view preconditions a need for continuous processing of additional information on top of meeting content. Therefore, we argue that compared to no self-view, the self-view may increase the cognitive load.

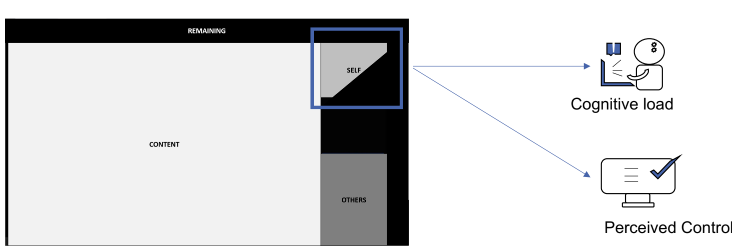


Fig. 1. The ambivalent nature of the self-view

Nature 2: Self-view and Control. Current research focuses on the self-view effect on self-awareness, particularly in relation to public self-awareness which is, based on our understanding, a pre-state to feeling control over how we present ourselves to others. Thereby, self-awareness theory is the core underlying theory that explains the self-view's impact on self-awareness [12, 24]. Self-awareness refers to the ability to think about oneself, how one looks, and what one does, experiences, or thinks. It is related to self-attention and self-presentation as a means of controlling how one is perceived by others [61]. Self-awareness can be dispositional or situational and can be impacted by external representations e.g., via mirrors or the mirrored self-image in video meetings [12].

Looking at control theory, it becomes clear that awareness is a crucial step towards control. For instance, looking at self-control, individuals need to be aware of themselves to be able to control their actions (see e.g., [3, 89]). Based on this understanding, we argue that self-view does have a controlling nature. By being able to see one's depiction sent to others, one is in consequence more aware of how others perceive oneself. Thus, one feels more in control about how one presents oneself and consequently more in control to alter their view of oneself. To date, studies however mostly focus on dyads and the impact on self-awareness in no shared screens and do not go one step beyond, focusing on control (e.g., [1, 39, 57, 65, 66, 75, 93]). However, findings from survey studies and interviews suggest this increased perceived control and related comfort [5]. We thus hypothesize that a displayed self-view serves as a control mechanism and increases the perceived control compared to no self-view.

Further Influencing Factors. Besides the pure existence of the stimuli, differences between individuals, teams, and contexts also impact how we interact with a certain technology and how we perceive it. This also applies for the self-view. Thereby, on an individual level, user characteristics in form of personality traits can impact us [44]. Related to self-view's impact, findings from psychology show that self-esteem, satisfaction with own appearance, and social interaction anxiety seem to impact [65, 75, 90]. On a team level, we are not yet aware of such findings, but a study from Shockley et al. [82] explores the impact of team constellations on self-presentation, however with cameras on versus off in total. Especially the closeness to others and hierarchical structures seem to make an impact, as people tend to care most about their self-presentation when impression-formation processes are involved. In general, the group size does influence communication and (gaze) behavior [63]. Similarly, the task also impacts as tasks requiring a high focus on the self, such as job interviews or self-presentations, may induce self-awareness based on design [42].

So overall we can see that both, the way the self-view is presented but also the individual, team, and context characteristics may impact the perception and interaction with the self-view and so far, not all factors are observed in the context of the self-view. Thus, we focus on a novel set-up with a shared screen in a triadic person setup and a neutral decision-making task.

2.3 Eye-tracking and Visual Attention

The user's attention is closely related to the individual's perception and the way the stimulus creates impact. Human attention is a complex process and can be distinguished in different forms of attention, amongst visual attention. When thinking of visual attention [74], it is important to understand that different forms of visual attention exist. Based on the eye-mind hypothesis, eye-tracking is thereby recognized as a useful means to investigate the so-called overt attention which covers the center of focus from a user [74].

In contrast, covert attention refers to attention not visually focused on and describes a state where the brain attends to an object without extrinsic behavior showing it. It is typically measured via brain signals, such as electro-encephalogram (EEG) [17, 21]. In HCI and the Information Systems (IS) community, a lot of studies focus on visual attention in e-learning or websites for usability improvements as well as understanding certain user behavior in interactive systems in general (see e.g., [15, 53, 60, 67, 86, 95]). In the context of video meetings, existing studies often leveraged eye-tracking data to understand the so-called social gaze. Predominantly they focus on joint attention of multiple persons in a meeting on the same topic and how it can be visualized (see e.g., [60, 79]). Besides, there also exist studies focusing on eye contact, also called mutual gaze [91, 92]. Concerning the gaze on the self-view, only a few studies exist. Vriends et al. [90] focus on gaze patterns of socially anxious versus not socially anxious participants on their self, including a discussion task with no shared screen and thus fewer stimuli. An additional study does solely briefly report that most fixations happened in the first minute [19]. Further, George et al. [33] report in a recent paper gaze patterns when showing participants recorded meetings. As an additional way to analyze gaze in video meetings, it is also possible to measure the overall attention of the user on the interface layout. Exemplary studies are those from Kuzminykh and colleagues aiming to classify attention types and functionalities [58] or studies that observe gaze patterns in creative meetings and video lectures [10, 88]. A focus on physical meetings and eye-tracking is also visible in the field of social psychology (see e.g., [11]).

When analyzing visual attention using eye-tracking, various measures exist. Fixations, described as eye movements focusing on a stationary object of interest, are the basis for analysis. Low-level measures such as fixation, fixation duration, fixation rate, fixation duration mean and number of fixations are traditional and common eye-tracking metrics [46]. Furthermore, area of interest (AOI)-based analysis is a popular means for stimuli-driven experiments. The screen is divided into areas containing similar semantic information, annotated as one AOI each. Based on each AOI, metrics can be calculated and compared. Scanpaths showing the sequence of fixations, share of gazes per AOI, number of fixations per AOI and gaze duration mean or median per AOI are most often used [46]. Beyond measurements spatially per AOI, fixation and saccade metrics can also be calculated temporally (see e.g., [83]) and show changes in fixations over time.

3 Method

We developed an experimental design that allows us to discover the diverging impact of the self-view feature's design throughout three consecutive and comparable meetings. In the following, we present our experimental procedure, the design variants of the self-view, the participant sample, and our data collection and analysis strategy.

3.1 Procedure

The experimental procedure during each session is presented in Figure 2 and consists of three overarching phases: (1) Onboarding phase, (2) Execution phase, and (3) Perception phase.

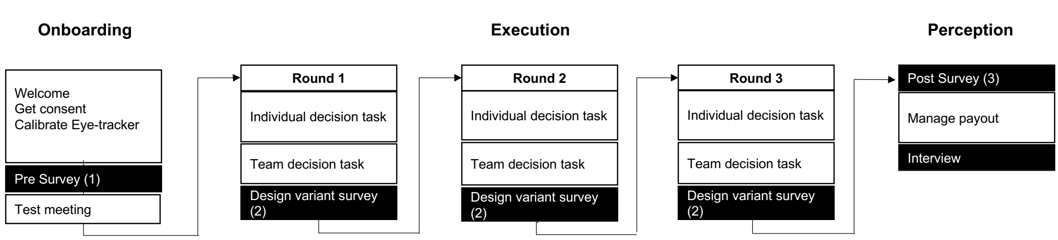


Fig. 2. Procedure

During the onboarding phase, users are asked for their consent and guided to the experimental rooms where the eye-tracking device is calibrated. The calibration is checked by a research assistant to ensure a good calibration fit. After reading the instructions, the participants fill out the pre survey to collect their current state of cognitive load. Afterward, they are introduced to a test meeting to familiarize with the software and the functionalities.

Afterwards, participants begin the execution phase which comprises three video meetings, each involving a decision-making task. As tasks, we chose the NASA survival on the moon task, and two alternative versions of the task which are Lost in the desert, and Lost at sea task [26, 36, 59]. The tasks were always presented in the same order. As we were interested in the impact of the design variants, we did not alter the task order. However, we conducted a post-hoc analysis of differences in task performance to check for learning effects. We did not find any learning effects in task performance between the rounds (S1: R1: M=39.31 (SD=4.27) MD=40.00 (IQR=4.00) vs. R2: M=57.23 (SD=11.12) MD=60.00 (IQR=12.00) vs. R3: M=65.38 (SD=10.18) MD=64.00 (IQR=22.00); for S2: R1: M=37.20 (SD=7.73) MD=36.00 (IQR=12.00) vs. R2: M=59.60 (SD=11.36) MD=62.00 (IQR=12.00) vs. R3: M=61.60 (SD=8.32) MD=62.00 (IQR=14.00), less points symbolize a better performance). Learning effects due to the VMS tool are already existing in the first round due to the preceding familarization meeting. Each round includes the following steps: The participants receive an instruction that informs them about the current scenario (i.e., survival on the moon, lost in the desert, lost at sea) and the task of ranking the items based on whether they would like to keep them or not (from 1: most important to 15: least important). Afterwards, participants provide their individual solution within three minutes. Then, they are directed to a meeting with two additional participants which lasts for 12 minutes after all participants have arrived in the meeting. In this meeting, participants discuss their individual solutions and are instructed to find a joint solution. After the meeting ends, the participants are navigated back to a survey where they are asked about our constructs of interest (cognitive load, perceived control, self-awareness) as well as the perception of how often they looked at themselves. This procedure is repeated for the second and third task. To discuss the items, in each meeting, one of the participants is sharing their screen. Every participant is sharing the screen in one meeting only. Thus, all participants are partaking in two meetings without sharing their own screen.

Finally, in the perceptions phase, the participants are asked to complete the final survey including control factors such as demographics, experience, or personality. Based on random selection, one of the meeting participants was selected to partake in an additional follow-up interview. Afterward, the participants get compensated and debriefed. The compensation consists of a fixed and a variable component to ensure that they engage in finding an optimal solution to the task. The variables component was chosen dependent on the correctness of the solution of the tasks. We received a data protection and ethics approval from the local institutional review board.

3.2 Design Variants

During the conducted study, participants engaged in three video meetings in a row in a team of three participants. Within the meetings they are exposed to an experimental stimulus, the self-view design variant. This stimulus is similar for all members of the team and over all rounds of meetings (i.e., between-subject design). We used the self-hosted open-source video meeting software JITSI (see https://jitsi.org/) which allows the manipulation of the self-view feature. In our study, we selected two different design variants: Design variant S1 constantly confronts users with their mirrored self-image. Design variant S2 does not offer a mirrored self-image. These design variants are visually depicted in Figure 3. We decided to choose a between-subjects design to observe the impact of the self-view in a row of meetings and by this be able to see changes over time.

3.3 Participants

We recruited 57 students from a local student participant pool at our university. We formed random teams of three participants each, including one female and two male participants, and assigned one design variant to the whole team. In total, our sample consists of nine teams (27 participants) assigned to design variant S1 (Self-view) and ten teams (30 participants) assigned to design variant S2 (No Self-view). The overall sample included 18 female and 39 male participants. To simulate a realistic team environment with mixed-gender teams, we designed the experiment sign-up process to ensure that each team consisted of two male and one female participant. 63% of the participants were aged between 21 and 25 years. 16% of participants were between 18 and 20 years and an additional 16% were between 26 to 30 years old. The remaining three participants were aged

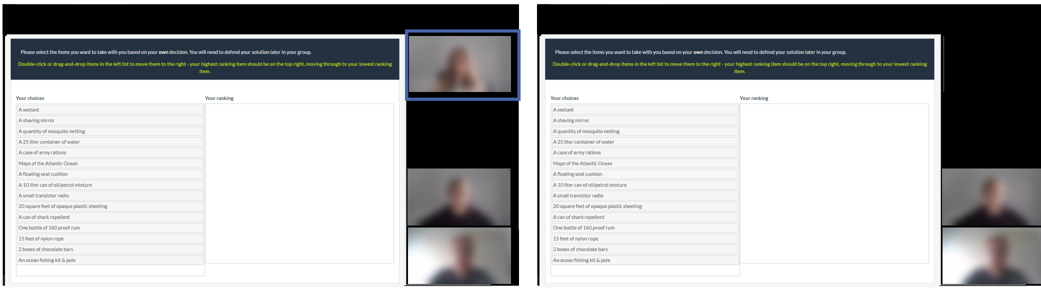


Fig. 3. Design Variants: Alternative with Self-view (S1) depicted in the left, Alternative without Self-view (S2) depicted in the right

between 31 and 35 years. All participants indicated to be very experienced in video meetings due to online lectures, group work, or leisure time activities (M=4.53 (SD=0.60) Median=5.00 (IQR=1.00)), equal for both design variants).

3.4 Data Collection

We utilized a comprehensive approach to collect data, incorporating quantitative measures through subjective questionnaires, qualitative insights from semi-structured interviews, and objective eye-tracking data. Eye-tracking provides continuous and detailed insights into participants' visual attention, unaffected by subjective biases present in the survey data. By this approach, we also provide a comparison of individual's perception and actual visual behavior. Furthermore, we captured the current window (i.e., survey, meeting software) using a Logitech webcam to record audio and video for analysis. We divide the analysis of results into three steps (i.e., perception, deep-dive, combination of both steps) and structure the data collection and results accordingly.

3.4.1 Perception.

Survey. Regarding the collected perception data, our experiment includes three types of surveys.: pre, design variant, and post survey. We used the pre survey (1) to get a baseline value for our time-dependent construct of cognitive load (modified Nasa TLX questionnaire [38]). To be able to monitor temporal changes between a series of meetings, a design variant survey (2) was issued after each meeting round and targeted our user state constructs. We asked for cognitive load (scale as above), perceived control (adapted from [77]), and momentary self-awareness (public, private, and situation, [34]). Moreover, we asked participants for their perceived frequency of looking at themselves as a manipulation check. In the post survey (3), we ask for control variables such as demographics (age, gender, eye-related diseases or vision control), previous experiences on task topics and video meetings, and personality-related user characteristics such as self-satisfaction with face/body appearance, self-consciousness [29], and Big-5 short scale [76]. An overview of all self-created or modified measurement items can be found in Table 5 in the Appendix.

To assess the perceptions collected in our survey, we analyzed the questionnaire items. As our constructs of interest are latent constructs, we assessed the outer factor loadings and Cronbach's alpha[87]. To ensure cutoff rates, we removed one item (item 4) from our less-established control construct. We provide an overview of all external factor loadings and cronbach alpha values in the Appendix, Table 5. In the next step, we calculated the mean value over all items for each construct and analyzed the data. As our data is not normally distributed, we use the non-parametric Wilcoxon rank-based test to observe between design variant effects and Friedman tests and pairwise Friedman tests to test effects between the rounds of meeting rounds using the "pgirmess" package in R [31].

Interview. As a complementary qualitative data source, we conducted semi-structured qualitative interviews with one randomly selected participant from each team subsequent to the experiment. In total, we performed interviews with 9 participants for design alternative 1 and 10 participants for design alternative 2. As we aimed to not bias the perception of individuals and aimed to have immediate feedback after the experiment with a comparable reflection period, we were only able to collect data from one team member each. To gain insights into qualitative perspectives from other participants as well, we asked them for feedback on experimental design and self-view use as well as self-view designs in a qualitative open-text question after the survey. Findings from these individuals are comparable to the interview data. The questions in the qualitative interview focused on the experiences with the conducted meetings, specifically with regard to the constructs of the survey as well as their general perception of the design variant. We also asked for changes over time. Besides, we asked about design improvements for future self-view designs.

The interviews lasted on average 20 minutes. We recorded, transcribed, and analyzed the interviews through a deductive thematic analysis [8], based on our constructs of interest as well as the topic of design. The first and second authors independently coded 20% of the interviews. Afterward, they discussed the codes and iteratively developed a coding scheme. Finally, the first author coded the remaining interviews. Our coding structure is visible in the Appendix, Table 7, 8, and 9.

3.4.2 *Deep Dive.*

Eye-tracking Data. To obtain continuous insights, we collected eye-tracking data. We used Tobii 4C eye trackers with 90hz sampling rate with a required research license and comparable screens having a resolution of 1920x1080 px for all meetings. Participants were seated on non-moveable chairs and instructed to move as little as possible to avoid problems with the accuracy of eye-tracking. For calibration, we used the default calibration measure provided by Tobii including six calibration points. For validating the calibration, we used a nine-point-based grid. We pre-processed the collected raw data with Python and transformed the eye-tracking data into fixations using the filter provided in the PyGazeAnalyser software package [18]. In a next step, we calculated various fixation-based metrics such as fixation count and fixation duration based on each area of interest (AOI) [23, 41]. We defined four AOIs based on semantic similar areas as depicted in Figure 4: self - containing the own image of the respective participant colored in medium grey, others – containing the image of the other participants in dark grey, content – containing the image of the shared screen when another participant was sharing the screen colored in light grey and remaining – containing further areas on the screen colored in black. We controlled for self and other extended AOIs due to possible calibration problems but did not find any needed adjustments of AOI sizes.

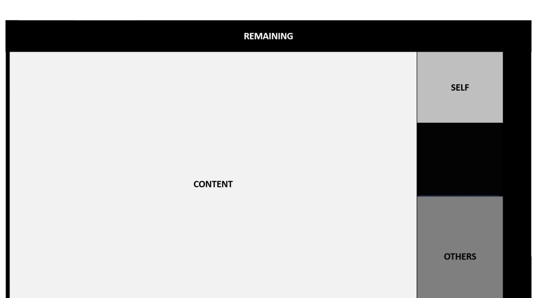


Fig. 4. AOIs

Table 1. Eye Metrics calculated with their description

Metric	Description
Share of fixation counts	Counts of fixations in an AOI divided by overall fixation counts
Share of overall fixation time	Time spend in an AOI divided by overall time spent
Median fixation duration	Median value of the fixation duration in an AOI
Mean fixation duration	Mean value of the fixation duration in an AOI

To ensure comparability, we respectively calculated the share of time spent on the AOI and the share of fixation counts spent on self-view. We further calculated fixation duration mean and median in milliseconds per AOI to see whether the fixations were just quick glances or longer looks at the AOI. An overview table including the metrics and their descriptions can be found in Table 1. We focused on the participants who experienced the self-view feature and used only recordings of participants who did not share their screen as a moderator. Due to data collection problems, we excluded data from two participants, resulting in an overall sample of data from 50 meetings. We present the findings in form of heatmaps, fixations per AOI over time, and transition matrices indicating changes between AOIs (as done in [37]).

3.4.3 *Combination of Both Perspectives.*

Identification of User Groups. To combine the insights from both perspectives, we apply clustering based on the perception data. As we do not want to define the number of clusters upfront, we explicitly voted for hierarchical clustering [51]. We scaled all variables, calculated distance matrices, and then conducted a thorough evaluation of different hierarchical clustering methods. To identify the best-performing algorithm and number of clusters, we used the following quality measures: cluster coefficient, elbow plot, silhouette width, Dunn index, and connectivity via the R Package NBClust. We ended up using ward.d. as clustering method, finding three clusters. The cluster coefficient has a value of 0.95 (scale 0 to 1, 1 indicating full explainability). Both, the elbow plot and NBClust majority vote recommend using three clusters for both approaches. The average silhouette width is 0.41. Dunn Index has a value of 0.2 and connectivity a value of 15. One explanation for the limited performance is the skewed data formation due to the share of fixation counts and the differences in especially increased control or cognitive load for some user groups that are smaller than the average group of users. We carefully examined whether these responses are due to outliers, but concluded that their gaze patterns are in line with their perception on frequently looking at self and the findings on high load or high/ low control are in line with interview findings.

4 Perception: Ambivalent Nature of Self-View Use in VMS (RQ 1)

In the first perspective, we focus on the two constructs related to the ambivalent nature, the cognitive load, and the control and underline our quantitative findings with interview quotes. We present the descriptive statistics in the following format within the text (mean value (standard deviation) median (interquartile range)). In addition, the summary statistics are depicted in Table 2. For our interviews, we refer to the participants by providing first their assigned design variant (S1: constant self-view, S2: no self-view) and afterward a generic participant ID. All codes and proof quotes are visible in Table 7 in the Appendix and highlighted in italics in the text.

4.1 Effect on Cognitive Load

The effect of the distracting nature of the self-view is instantiated via the construct of cognitive load. We compare the cognitive load on an overall level and observe temporal changes, similar to the identified thematic coding.

Level of Cognitive Load: The comparison of cognitive load between the design variants is visible in Figure 5. When looking at the evolvement of cognitive load between the design variants (Figure 5 on the left), we can see a slight increase in cognitive load in design variant S1, constant self-view, compared to no self-view (overall S1: M=4.38 (SD=1.28) MD=3.91 (IQR=1.83), S2: M=3.75 (SD=1.02) MD=3.33 (IQR=1.50)), significant overall effect with W=4743, p=0.001). We further compared the significance of the effect for each round and identified that in Round 1 (S1: M=4.18 (SD=1.28) MD=3.83 (IQR=1.17), S2: M=3.66 (SD=0.82) MD=3.67 (IQR=1.00), W=498.5, p=0.136) and Round 2 (S1: M=4.34 (SD=1.30) MD=4.00 (IQR=1.67), S2: M=3.74 (SD=0.87) MD=3.42 (IQR=1.17), W=504.5, p=0.112) no significant difference is found. In Round 3 we observe a significant difference between the design variants (S1: M=4.62 (SD=1.27) MD=4.33 (IQR=1.08), S2: M=3.84 (SD=1.03) MD=3.67 (IQR=1.46), W=564.5, p=0.011). Overall, the level of the cognitive load was perceived as moderate, in line with the findings from the qualitative interviews where 7 out of 9 interviewed participants for design variant S1 and 10 out of 10 participants for design variant S2 said that they have a *low or moderate* level of cognitive load.

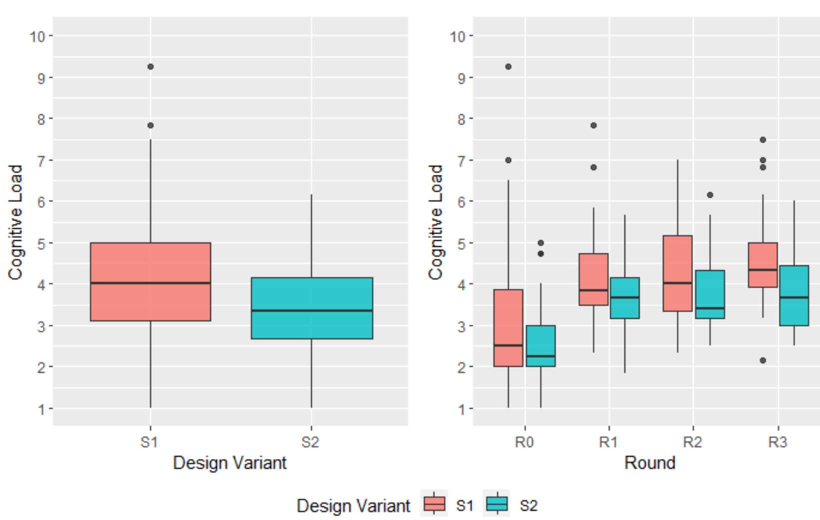


Fig. 5. Descriptive comparison of cognitive load between the design variants over all rounds (left) and per round (right)

Temporal Changes: Besides the overall perceptions per design variant, we are also interested in changes over time, especially within a design variant. We identify that cognitive load is rather stable over time in both design variants (S1: R1: M=4.18 (SD=1.28) MD=3.83 (IQR=1.17) vs. R2: M=4.34 (SD=1.30) MD=4.00 (IQR=1.67) vs. R3: M=4.62 (SD=1.27) MD=4.33 (IQR=1.08); no significant difference in Friedman test: Chi-squared =4.8911, p=0.086; for S2: R1: M=3.66 (SD=0.82) MD=3.67 (IQR=1.00) vs. R2: M=3.74 (SD=0.87) MD=3.42 (IQR=1.17) vs. R3: M=3.84 (SD=1.03) MD=3.67 (IQR=1.46), no significant difference in Friedman test: Chi-squared =0.9333, p=0.6271, except a significant increase compared to pre-survey baseline value base S1: 3.33 (1.98) MD=2.34 (IQR=1.50), Chi-squared =26.287, p=0.000; S2: 2.53 (1.09) MD=2.25 (IQR=1.00), Chi-squared =36.895, p=0.000). This can again be confirmed with the findings from interviews where 12 out of 19 participants stated a *stable* level of cognitive load. Besides, 7 out of 19 interviewed participants perceived changes in cognitive load. Most thereby stated slightly *increasing* cognitive load toward the end. However, this seems to be rather a fatiguing effect than pure cognitive load as they indicated that *"it was just,[...] because of the tiredness that came up"* (S2:a22). In direct contrast, one participant reported a *decrease* in cognitive load due to getting used to the task and participants. Further, two participants reported increased cognitive loads associated with their role in the meeting. However, it remains questionable whether the moderation itself or the task led to the cognitive load as they highlight that *"[they] don't think it's because I shared it, but because it was an exercise that we had more thoughts about"* (S1:a18). To conclude, while most participants recall the tasks to be equally difficult, minor exemptions to recall fatiguing effects that make it more difficult or starting problems making the first task more difficult exist.

4.2 Effect on Control

Symbolizing the control mechanism of the self-view, we investigate perceived control. We compare the control on an overall level as well as identified rationales for being in or out of control, and observe temporal changes, similar to the identified thematic coding.

Level of Control: For control, we see slight changes as depicted in Figure 6 which are marginally not significant (overall S1: M=4.70 (SD=1.30) MD=5.00 (IQR=1.67), S2: M=4.29 (SD=1.32) MD=4.33 (IQR=2.33), no significant effect, W=4270.5, p=0.052). In the interviews, three participants explicitly stated feeling a *high* level of control when being confronted with the self-view.

Rationale for Control: Taking a deeper look at what led to the control when being confronted with the self-view, five participants said that the pure existence of the *self-view picture* of themselves helped to get control. Besides, however, one participant mentioned *other* mechanisms to gain control as the other's image of oneself is always only partially influenceable. Besides, especially for the group confronted with no self-view, the perceived control seems to vary drastically within the design variant (IQR for S2: 1.67 in Round 1, 1.92 in Round 2, 2.33 in Round 3). This can be supported by our interview findings showing two differences of perceived control. First, individuals who perceive no control, and second individuals who perceive control because they compensate with thoughts of self. For the first group, 4 participants perceived *low* control due to the missing self-view and missed control with *no coping* mechanisms applied. In total, 4 out of ten interviewees indicated missing control and stated that they would have preferred a self-view. Thereby, some participants felt irritated, like losing themselves as visible by a quote from an interviewee stating that *"I lost myself as an image, so to speak, that I was now sitting there quite normally and never thought to myself, ok, maybe I have to sit up a bit straighter [...]"* (S2:a34). However, showcasing the second group with *high* perceived control despite the missing self-view, 5 out of 10 interviewees reported that they applied *coping* mechanisms, and thus felt more in control by thinking more about themselves and how they could be perceived by others.

Temporal Changes: In terms of perceived control over time, we see slight decreases in design variant S2, no self-view, and a rather stable development in S1 (Round 1: M=4.69 (SD=1.20) MD=5.00 (IQR=1.25) vs. Round 2 M=4.80 (SD=1.39) MD=5.17 (IQR=1.58) vs Round 3 M=4.62 (SD=1.32) MD=4.83 (IQR=1.58) in S1, no significant overall effect: Friedman Chi-squared = 5.3043, p=0.0705). In design variant 2, no self-view, the control is highest in round one, and descriptive differences between the rounds are visible (Round 1 M=4.67 (SD=1.25) MD=4.67 (IQR=1.67) vs. Round 2 M=4.08 (SD=1.27) MD=4.17 (IQR=1.91) vs. Round 3 M=4.12 (SD=1.45) MD=4.33 (IQR=2.33) in S2, no significant change in Friedman test Chi-squared = 4.4706, p=0.107). Looking at the qualitative findings, this slight decrease might be because these coping mechanisms are most important in first meetings when not knowing the others as participants stated to *"made sure that the camera*

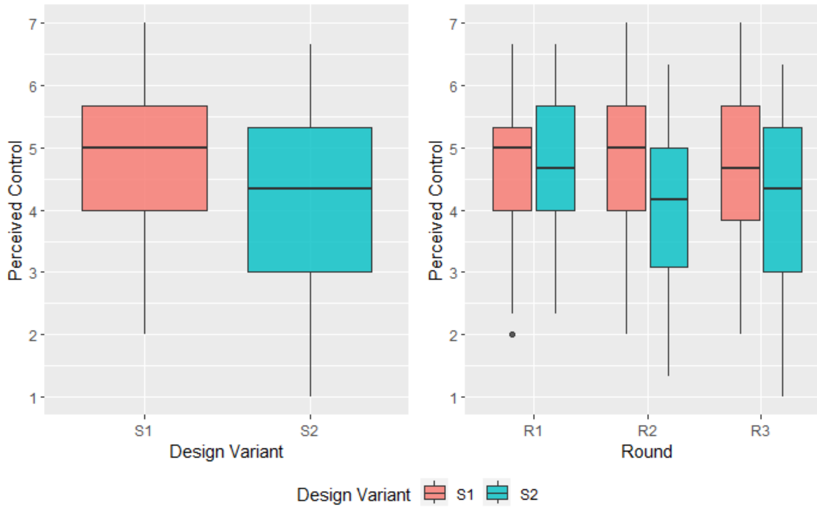


Fig. 6. Descriptive comparison of control between the design variants over all rounds (left) and per round (right)

Table 2. Summary statistics for cognitive load and control per round

Construct	Variant	Round 1		Round 2		Round 3	
		Mean (SD)	Median (IQR)	Mean (SD)	Median	Mean (SD)	Median (IQR)
Cognitive Load	S1	4.2 (1.28)	3.83 (1.17)	4.34 (1.30)	4.00 (1.67)	4.62 (1.27)	4.33 (1.08)
	S2	3.66 (0.82)	3.67 (1.0)	3.74 (0.87)	3.42 (1.17)	3.84 (1.03)	3.67 (1.46)
Perceived Control	S1	4.48 (0.89)	5.0 (1.25)	4.60 (1.05)	5.17 (1.58)	4.38 (0.92)	4.83 (1.58)
	S2	4.48 (0.85)	4.67 (1.67)	4.00 (0.91)	4.17 (1.91)	4.03 (0.92)	4.33 (2.33)

wasn't standing in any inappropriate or unflattering way in general. But at the beginning, it was just a matter of ten seconds. And after that, I didn't really pay any attention to it" (S2:a16).

5 Deep Dive: Visual Behavior During Self-view Use (RQ 2)

To get a deeper understanding on how users perceive the presence of the self-view feature, we analyzed the interaction with the self-view feature via eye-tracking technology and compared it to our findings from the interview. Codes and proof quotes are available in Table 8 in the Appendix and marked in italics in the text.

5.1 Overall Visual Behavior

The general visual behavior is depicted in Figure 7. In both design variants the share of fixation counts is highest for the content AOI (S1: M=76.50 % (SD=9.62%) MD=75.20% (IQR=14.62%), S2: M=77.96% (SD=8.51%) MD=78.97% (IQR=10.40%), followed by the AOI of the other participants (S1: M=16.76 % (SD=9.01%) MD=16.76% (IQR=12.24%), S2: M=14.11% (SD=7.48%) MD=12.57% (IQR=11.36%)). The self-view shows the lowest share of fixation counts (S1: M=2.63 % (SD=3.66%) MD=1.08% (IQR=2.27%), S2: M=0.19 % (SD=0.57%) MD=00.05% (IQR=0.14%)), however, participants spent approximately 3.18% of meeting time and 2.68% fixation counts on the self-view with a maximum share of time of 22.02% and share of counts of 16.67%. Similar patterns are visible for the overall time spent on the AOI. A similar perception has been stated in the interviews where 78 percent of all interviewees stated that they predominantly recall focus on the *content*.

Individuals confronted with a self-view have a slightly higher share of fixation counts and fixation time on people in general, compared to those not confronted with self-view (S1 (others + self): M=21.16% (SD=15.74%) MD=17.84% (IQR=14.51%), S2 (others + self): M=16.91% (SD=10.38%) MD=12.62% (IQR=3.80%)). This can be underlined with statements from the interview indicating that participants feel more immersed on the content when no self-view is there, and *"pretty much sink into the task."* (S2:a19).

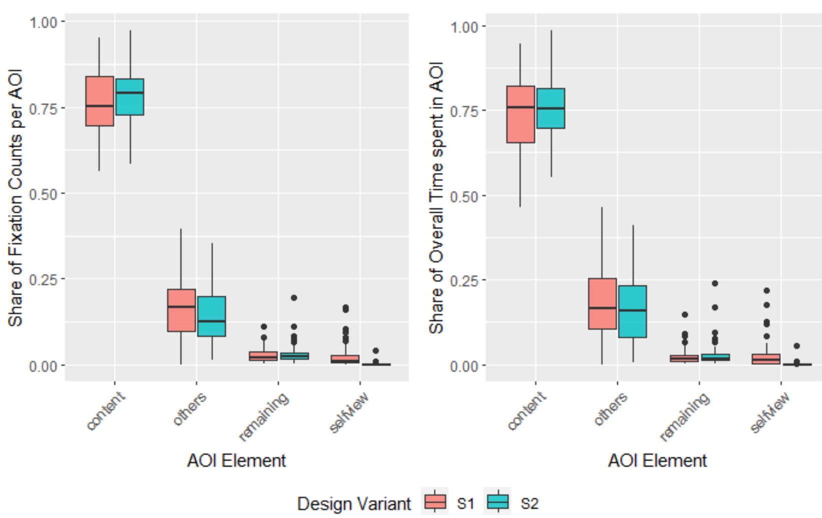


Fig. 7. Share of fixation counts (left) and overall time spent per AOI (right) for both design variants

5.2 Visual Behavior on the Self-view

The fixations on self-view show a broad range with the share of fixation counts and the share of overall fixation time differing between 0 and 16 % in design variant 1. This is also visible in our interview findings where participants indicated a range from looking *often* at self, stating *"as soon as you had the time or the opportunity to look at, [...] you tend to do that"* (S1:a21) to others saying to *not look* at the self by actively avoid looking at one image. As the total time and counts of fixation spent in the AOI have a wide variance, we want to better understand what in-depth patterns are visible with regard to this self-viewing behavior. Therefore, we take a look at the differences in intensity to look at self via heatmaps, the timing when to look at self, and subsequent and previous AOIs when looking at self. An overview of the summary statistics related to self-viewing overall is visible in Table 10 in the Appendix.

Heatmaps. We created heatmaps for all participants per meeting and analyzed them based on the intensity of the coloring. The warmer the color gets, the more the focus is set on the specific area. We see vast differences when it comes to the intensity of looking at the self-view. Three main distinctions can be made: First, some participants do not look or only very limitedly look at themselves even though being presented with their self-view on shared screens. This does contradict the indicated reflexive nature of the self-view and the attention-grapping nature of people representations, at least for the picture of self. In total, 18 participants meetings from our sample show such behavior. An example is visible in left part of Figure 8. Statements from the interviews underlining this behavior is a high focus on content whereas *low focus on self* during productive meeting phases (S1:a14) or active avoidance of the self-view.

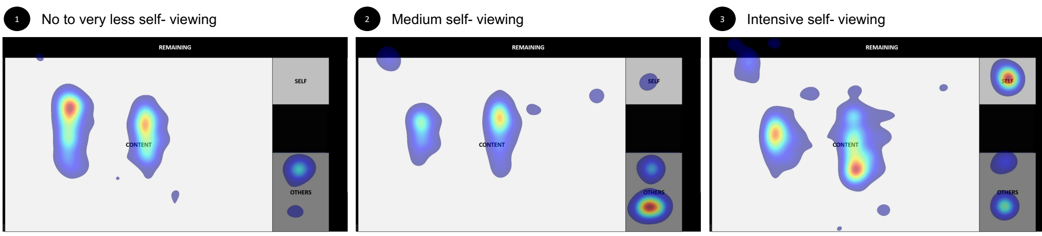


Fig. 8. Comparison of heatmaps based on the intensity of self-view behavior

Second, there is a broad range of participants who focus only *limited* on themselves. This is visible in Figure 8 in the middle. In total, 21 participants' meetings belong to this in our sample. These participants seem to be looking at themselves during the meeting, however, do not look at themselves very intensively. Interview statements from participants in this group highlight that they occasionally look at themselves but still have a high focus on the content and/or others. As participant S1:a31 stated that *"if (you) start focusing more on the meeting, (you) will just forget"*.

Third, there seem to be some participants that *intensively* look at themselves. An example can be found in Figure 8 on the right. In total, 11 participants meetings belong to this group. Based on the interviews, these participants are the ones that look at themselves for example when making a statement, check whether the camera setting is still ok, or seek control due to further reasons.

Self-view over Time. We also analyze where participants look at over time. This is motivated by the fact that participants indicated that they perceived to have differed their self-viewing behavior depending on how focused they were, as for instance stated by S1:a14 who stated to forget looking at themselves in productive phases or recalled to mostly focus on themselves *"mainly during waiting periods"* (S1:a14) which happened either in the beginning or towards the end. Therefore, we calculated time stamps based on the audio- and video-recording when they were focused (i.e., until all items were sorted), mildly focused (i.e., after first sorting of all items until they ended discussing the items), and talking (i.e., until the end of the meeting if time remained not talking on the items). We included separation lines in the plots to ease the analysis (dotted red line: end of focus, first black line: end of mild focus, second black line: end of talking).

Over time, there seem to be three patterns: First, *constantly* looking at oneself throughout the whole meeting without a distinguishable pattern. This is by far the most prominent group. In total 34 participant meetings belong to it. Within the group, there is a difference between participants looking often and participants not looking often at self, as it can be assumed based on the introduced heatmaps. It is visible in Figure 9, number 1. Our interviews, for instance underline this with statements on the reflexive nature of the self-view or such as, *"as soon as you had the time or the opportunity to look at, you tend to do that"* (S1:a21).

Second, looking at self towards the first part of the meeting, when being *focused* on sorting the elements. This is the smallest group containing seven participant meetings. It is visible in Figure 9, number 2. In the interviews, this behavior was not recognized and stands in contrast to some statements of individuals that *"will just forget (about the self-view), especially if I start focusing more on the meeting"* (S1:a31). One possible reason could however be that individuals confirmed how they physically present themselves and look at their image when being insecure about the image

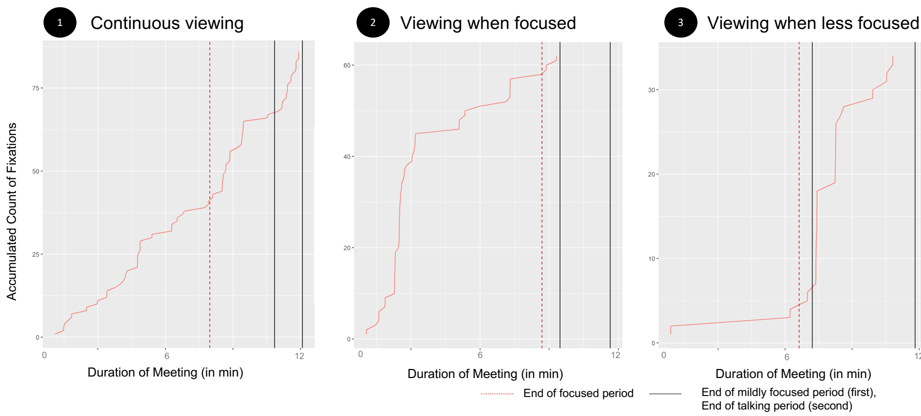


Fig. 9. Comparison of individual patterns showing Accumulated Counts on Self-view

they currently send to others. This is visible from the following statement *"you do not know the participants and you don't want them to go out there and think, oh, she was a bit of a handful (...) That is why I thought about it. I think it became less, (...) then I didn't have the doubts anymore"* (S1:a14).

Third, looking at self more often towards the end, in *no or only mildly focused* work. This group contains nine participant meetings. It is visible in Figure 9, number 3. In our interview, participants indicated that they are more looking at self towards the end, or when less concentrated. A possible explanation for such a behavior could be that participants look more often at self when they are part of the topic of the talk or get bored.

Transitions Between AOIs. In a final step, we identify switches between AOIs via transition matrices. We explicitly focused on areas visited prior or after looking at self. Overall, we see that most transitions toward the self-view stem from the content. Transitions out of the self-view mostly go to content or other participants. As these AOIs are the ones containing the most information relevant to solving the task and interacting within the VMS, this is natural. When looking at participants separately, we again see differences as visible in Figure 10.

A first pattern is visible for six participants in Figure 10, number 1 and 2, who tend to visit the self-view solely after looking at the content and also only proceeding back to the content, or sometimes others, afterward. This pattern may relate to participants who stated to be very much focused on the content. Consequently, they are also only able to transition to self from this AOI. A second pattern covers the opposite and describes eight participants first looking at themselves after the content or others and then going back to others. This pattern is visible in Figure 10, number 3, and belongs to participants describing that they checked whether others find their own ideas strange or not and sought confirmation after making a statement. Similarly, in twenty participant meetings depicted in Figure 10, number 4, we see that a high number of participants also visit the self-view from remaining AOIs. After the self-view participants differ again depending on their subsequently visited AOI, as visible when comparing both number 4 plots in Figure 10. Finally, no clear pattern that differs from looking at any relevant AOI and going back to any of them is visible for 16 participant meeting plots depicted in Figure 10, number 5. Thereby, plots in this group may either contain multiple types of transition groups outlined above or belong to individuals that are just purely reflexively looking at themselves when they had the opportunity.

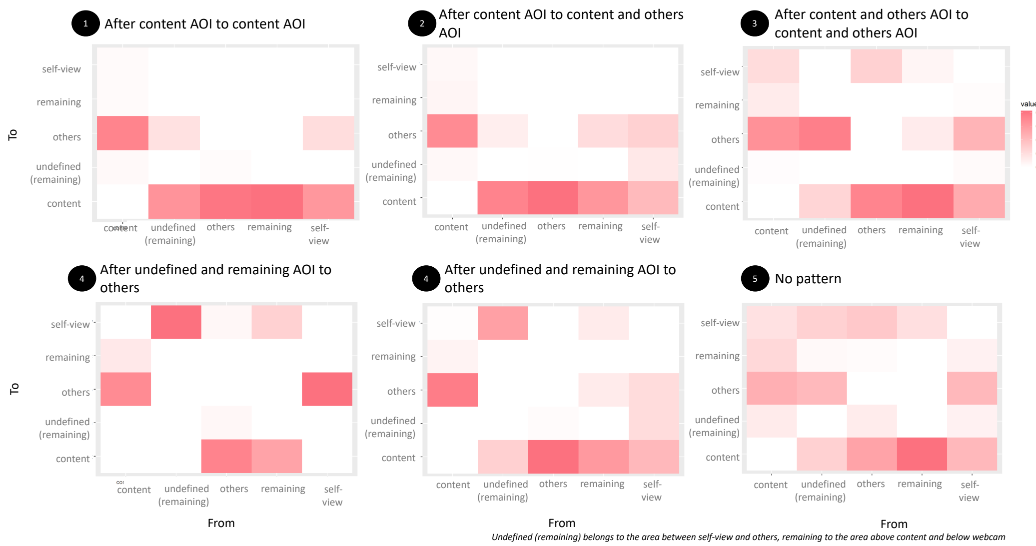


Fig. 10. Comparison of transition matrices between different patterns

6 Combination of Both Perspectives: Clustering of User Perceptions

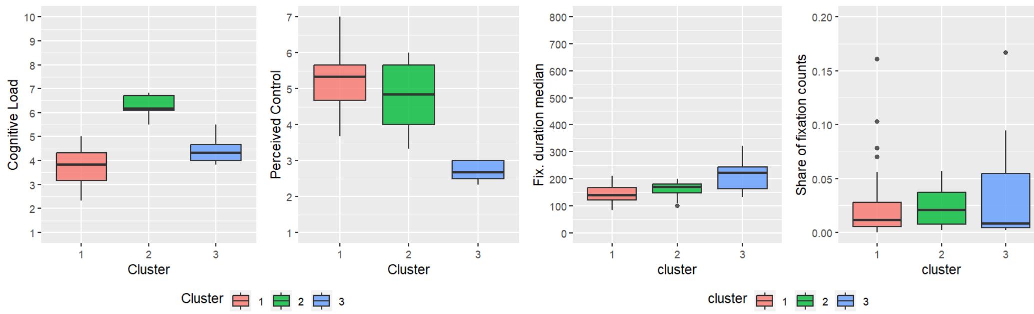
To derive a deeper understanding of how the observed differences in the users' perception relate to visual behavior (i.e. share of fixation counts and median fixation durations in the heatmaps in Figure 8) and user characteristics, we perform a cluster analysis [51].

Clustering. We conducted a hierarchical clustering analysis and present the clusters in Figure 11a and Table 3. We identify three clusters based on perceived control and cognitive load. Cluster 1, called "Benefiting Users" depicted in red, shows a comparably low cognitive load and high perceived control. This cluster covers individuals that seem to benefit from the positive nature of the self-view and have high control without too much costs of induced cognitive load. Cluster 2, called "Cognitively Challenged Users" depicted in green, covers participants indicating high cognitive load and a medium level of control. This resembles the cognitive load inducing negative nature of the self-view, especially as the reported perceived control is slightly decreased compared to those in cluster 1. Cluster 3, called "Control Losing Users" depicted in blue covers participants having a low cognitive load and low level of control. This resembles individuals that do not benefit from the self-view and still perceive limited control.

Table 3. Description of cluster differences (Mean (SD), Median (IQR))

Cluster name	Cognitive load	Control	Median Fix. dur. (in ms)	Share of fix. counts (%)	Size
1: Benefiting Users	3.79 (0.76)	5.27 (0.83)	138.63 (39.43)	0.02 (0.03)	35
	3.83 (1.17)	5.33 (1.00)	133.24 (44.41)	0.01 (0.02)	35
2: Cognitively Challenged Users	6.27 (0.48)	4.79 (0.96)	159.61 (35.70)	0.03 (0.02)	8
	6.17 (0.63)	4.83 (1.67)	169.32 (31.92)	0.02 (0.03)	8
3: Control Losing Users	4.43 (0.61)	2.71 (0.30)	213.32 (67.37)	0.04 (0.06)	7
	4.33 (0.67)	2.66 (0.50)	222.06 (80.45)	0.01 (0.05)	7

Cluster Group Characteristics and Effects. In a next step, we report the differences in visual behavior between the clusters, visible in Figure 11b and Table 3. Cluster 1 covers individuals that



(a) Cognitive load and control

(b) Fix. med. duration and share of fix. counts

Fig. 11. Comparison of cognitive load (left) and perceived control (right) (a) and Comparison of fix. med. duration (left) and share of fix. counts (right) (b)

seem to have minimal interaction and look at themselves for the least duration per fixation and with a comparably low fixation count. Cluster 2 individuals look at themselves with an increased interaction, meaning a slightly longer fixation duration and show a high diversity in terms of fixation frequency. Cluster 3 individuals interact most intensively, spending a high fixation duration per visit. Descriptively, they look most frequently look at themselves, again however a very diverse pattern of share of fixation counts is visible. Thus, it seems like the duration of looking at the self is more important than the number of visits to the self-view.

To better understand where the differences in perception may stem from, we briefly take a look at user characteristics of the individuals. We follow the findings from Potthoff and colleagues and Barnier and colleagues [7, 75] and focus on self-satisfaction. In addition, we observe neuroticism and social interaction anxiety which may be drivers for a need of control or feeling of loosened control in social situations [65, 90]. Looking at the distribution of these user characteristics, we see differences between all clusters, especially for neuroticism and social anxiety. We see that the "Benefiting Users" (C1) Cluster descriptively seems to be least neurotic (C1: M=2.45 (SD=0.93) MD=2.00 (IQR=1.00), C2: M=3.06 (SD=0.94) MD=3.00 (IQR=1.25), C3: M=3.21 (SD=1.22) MD=3.50 (IQR=1.50)). In addition, the "Control Losing Users" (C3) descriptively have the highest rating of social anxiety and lowest self-satisfaction with their appearance (Social anxiety: C1: M=2.99 (SD=0.67) MD=3.20 (IQR=1.00), C2: M=3.40 (SD=0.88) MD=3.30 (IQR=0.95), C3: M=3.66 (SD=0.38) MD=3.40 (IQR=0.40), Self-satisfaction: C1: M=3.56 (SD=0.74) MD=3.50 (IQR=1.00), C2: M=3.81 (SD=1.16) MD=4.00 (IQR=2.50), C3: M=3.29 (SD=0.86) MD=3.00 (IQR=0.75)).

7 Perception of Self-View Designs and Future Design Improvements

Next to assessing the effects of the different design variants, we aim to gain an overall understanding of participants' perceptions of the investigated self-view design variant. Furthermore, we collected ideas for future design improvements from the participants.

7.1 General Perception of the Self-view Design Variants

All participants recognized their assigned design variant. We also learned that the general perception of the corresponding design variants differed. We organize these insights around the themes of perception of the self-view or no self-view, the codes are visible in Table 9 in the Appendix and highlighted in italic in the text.

Perception of the design variant with self-view: We identified that all interviewed participants being confronted with the self-view were used to a self-view from their experience with past video meetings. The overall perception of the self-view per se was rather positive, despite two participants indicating not liking to see themselves much. The first participant mentioned feeling *neutral* about the self-view. Another participant mentioned a *negative* perception and that they prefer no self-view and even avoided looking at themselves. Except for these two interviewed participants, all remaining participants indicated a *positive* perception and prefer having a self-view. They caught themselves looking occasionally and compared it to a reflexive behavior.

Perception of the design variant with no self-view: For participants not confronted with a self-view, we observed almost equally positive and negative perceptions of the design variant. From the ten interviews in the design variant, we observed that 50 percent did find the non-existence of the self-view *positive*. These participants mentioned being more focused on the meeting. According to one participant, that even ended with themselves feeling *"completely dissociated from my own appearance"* (S2:a32).

From the remaining 50% of interviewees, one participant felt *neutral* about the missing self-view, finding it uncommon but no problem. The remaining 40% of interviewed participants said they find it *negative*, irritating or uncomfortable at first.

7.2 Ideas for Self-View Design Improvements

We further asked participants about alternative design ideas for an adaptive self-view. The results are analyzed around the three dimensions of general attitude towards an adaptive self-view, the adaptation mechanism, and the adaptation visualization [27]. We chose these themes as they represent all design areas that can be tackled in alternative self-view designs. Codes are visible in Table 9 in the Appendix and highlighted in italics.

General attitude: Participants mostly showed a *positive* perception and agreed that an adaptive self-view would be a good idea. Overall, ten of our participants proactively mentioned that they would prefer a more flexible design of the self-view. Only three participants indicated *negative* attitudes and indicated that they prefer existing self-view designs and do not want other options.

Adaptation Mechanism: As a general question regarding adaptation, we asked participants to state possible mechanisms they would want to trigger an adaptive self-view. Overall, only few participants preferred user-driven *adaptable* versions such as a switch on/off version of the self-view. In favor of more adaptive approaches, system-driven adaptive mechanisms could continuously collect information from the user and the current meeting. As a first idea, *eye-tracking data* was proposed. Beyond, also triggers based on the specific meeting situation were proposed, including *task-based* approaches. One idea centered around differing the self-view presentation based on focused versus less focused situations. Another idea was to include information on how much one is related to the conservation topic, as some indicated they would like to have it while presenting, while others don't. As a last idea, *participant-dependent* adaptation mechanisms were proposed, especially on how well you know other participants, as participants also indicated to be interested in their depiction with friends being present.

Adaptation Visualization: In addition, also the visualization of the self-view could be altered.

Only one of the participants explicitly mentioned *filtering* techniques such as compression methods. Other participants indicated that they prefer the current visualization option already.

8 Discussion

In this section, we discuss our results, outline implications for design and the limitations of our work as well as future research opportunities. We provide a summary in Table 4.

Table 4. Summary of results for all perspectives

	Perspective	Quantitative Summary	Qualitative Summary
Perception	Cognitive Load (RQ 1)	Self-view increases cognitive load, an overall significant effect is visible Increased cognitive load effect is stable between meetings	Interview statements support quantitative findings Perception of tasks as equally difficult
	Control (RQ 1)	Perceived control is descriptively increased when self-view present, borderline significant Effect visible starting with second round When no self-view, control descriptively decreases over time	Interview statements support quantitative findings No self-view: mixed perception exist that explain the high deviations: either participants indicate missing control, or having control due to coping mechanisms
Deep Dive	Visual Behavior overall (RQ 2)	Most fixations are on the content or others Slightly higher people orientation for participants with self-view	Most participants recall focusing most on the content AOI No self-view: perceived higher focus and immersion than existing self-view
	Visual Behavior Self-view (RQ 2)	Self-view fixation count and median duration differ within the design variant Heatmaps: three types of users seem to exist (no to few intense, medium intensive, and intensive viewing) Over time: three patterns are visible (constant use, mostly when working in focus, when discussing and not working focused) Transitions: different patterns when it comes to AOIs visited before and after self-view Most people focus on non-described AOIs, followed by content or others	Findings on when and how often looking at self differ and confirm the findings from the visual behavior w.r.t heatmap groups, self-view over time Especially for no focused work participants recall looking more on others and self No findings for transitions from the interviews
Combination	User Groups (RQ 2)	Three user groups in alignment to visual behavior analysis and differences in perception are identifiable	Increased interaction, esp. fixation duration, with self-view seems to be related to negative effects Viewers encountering negative effects of control show descriptively lowest satisfaction with appearance, most neuroticism and social anxiety

8.1 Summary and Discussion of Study Results

We organize the results of our study based on our core findings.

1. The self-view increases cognitive load compared to no self-view

Looking at the perception data, we see that the ambivalent nature of the self-view is descriptively visible in our results. Our findings suggest that cognitive load significantly increases in follow-up meetings when being confronted with the self-view and an overall significant difference is visible. This may support the findings that the self-view indeed does increase cognitive load, even when visualized with a small size [42]. Thus, we contribute to the understanding of the diverging findings in the existing literature. Our study confirms the findings that the self-view is a driver for cognitive load and hence contributes to virtual meeting fatigue over time, which is increasingly researched [78]. Thereby, our results interestingly show the self-view's negative impact in a shared screen setting compared to existing studies with no shared screen (e.g., [42]). In light of findings from Miller and colleagues [65] which report significant effects depending on the size of the self-view, the small effect may be explained by the rather small size of our self-view, due to the shared screen.

2. The self-view seems to increase control compared to no self-view

Besides increased cognitive load, descriptive differences which are borderline significant (p=0.052) are visible for perceived control suggesting increased control of users when being confronted with the self-view. The quantitative difference can be underlined with our qualitative findings. One reason for being only borderline significant may again be the small size coupled with a task that does not focus on presenting oneself, as it has been done in other studies (e.g., [42]). Looking at data over time, we especially see a difference in the perceived control in follow-up meetings, suggesting that over time the control remains stable for an existing self-view but seems to slightly decrease for no existing self-view, as also visible in the qualitative findings. This finding supports the controlling nature of the self-view mentioned in survey studies [5] and public media [20]. With our study, we present first empirically grounded results for this effect. However, we also suggest that further studies are needed to better understand this positive beneficial effect of the self-view (e.g. comparing shared / no shared screen setting, temporal effect).

3. Different strategies to regain control exist with no self-view

In addition, the missing self-view feature seems to have an ambivalent effect: When not experiencing the self-view, our participants seem to either forget about themselves or even worry more about how they are perceived. This finding may explain how people react when being confronted with a video-mediated communication that is closer to naturalistic settings that also do not show a self-mirror but does lack a commodity feature in the video-mediated context they know. Thinking of future versions in remote collaboration, this may also lead to inspirations for the design of meeting spaces in the future. Regarding the missing self-view feature effect, we argue that one explanation for this may stem from coping mechanisms applied when feeling out of control, as also visible in the interviews [30, 89]. Participants lacking control applied a coping mechanism that leads to higher perceived control. We argue that this also shifts their focus to themselves and thus also leads to similar distracting effects as the self-view is said to normally have [30, 89]. We see that these coping mechanisms might increase the perceived control at first, however, in the series of meetings, this effect seems to reduce based on our descriptive findings. One reason may be impression formation processes that happen when first working together and a related effort spent on trying to give a good impression [61] which in turn leads to the increased focus on self and associated coping mechanism. However, this effect of the missing self-view is counterintuitive based on existing theories. We suggest that a habituation effect to the self-view as a common feature nowadays may serve as an explanation and urge researchers to explore this further.

4. Different visual attention strategies exist for overall self-view interaction, time-dependent interaction and transition from and to the self-view

In terms of visual behavior, we identify a strong focus on content for both design variants with over 70% of all fixations being spent in the content AOI and a high rate of fixations on the screen. Looking at self-view interaction per se, we identified a wide range of behaviors, ranging from not looking at self to looking at self for approximately 16% of the overall time. Over time, we see three patterns on how to interact with the self-view, either a continuous pattern, looking mostly at self when being focused or mostly looking at self in discussions. Finally, we observe differences related to AOIs visited prior to and after the self-view. While we do not see a clear pattern for most participants, some preferences to look at content or look at others prior to the self-view exist.

As current studies on visual behavior did not explicitly outline the self-view tile as a distinct element, we extend their "partner" AOI by splitting it into "others" and "self". We see that other participants are allocated with more fixations than the self and that the focus on content is still predominant, supporting findings of content focus in contrast to people focus in meetings having a shared screen [60, 91, 92]. Besides, in contrast to [33] who identified that users spent one-third of their time looking at something other than the screen in small discussion groups, we identified a high rate of fixations on the screen. This is in line with Brucks and Levav [10] who identified that participants looked less around and more at other participants in virtual meetings compared to physical meetings. One explanation for the different findings can be the shared screen and the incentivized task, underlining that the meeting context, especially task and interface layout, impacts the visual behavior and potential effect of technological stimuli in the meeting.

Looking at the self-view interaction per se, the identified individual differences in fixation on the self-view are in line with Vriends et al. [90] in a video setting and findings from a physical setting, where studies try to induce a mirrored image of oneself via a mirror or self-reflection when walking next to shopping windows (e.g., [14, 35, 62, 75]). We extend them with findings on the transitions and an overview of time-dependent differences. Our identified temporal changes of self-view interaction are in contrast to the study of De Vasconcelos Filho et al. [19] who identified that most fixations happened during the first minute. This may first be explained by the fact that the shared screen covers more area on the screen than the participants' AOI and is the most prominent AOI. Hence, participants may first fixate on the shared screen even though faces in general, especially one's own face, are very prominent stimuli [49]. Second, the self-view is no longer novel but more habituational in our study and research shows that more novel stimuli also attract gaze [43]. In conclusion, this means that individuals seem to interact with the self-view differently now compared to earlier studies. They no longer only use the self-view as a tool to quickly look at in the beginning and avoid it later on but use it throughout the whole meeting, resembling more the behavior shown in studies with mirrors or self-reflections in daily lives. Regarding transitions, we did not find prior work to compare the findings with. However, the different identified patterns can be explained. Since more than 70% of fixations are on the content it is only natural that people visit the self-view from this area, especially when they are highly focused. A more interesting pattern depicts the look at the self either from remaining AOIs showing no information but close to the webcam or from areas between self and others. Looking at self before or after looking into the camera can be explained by the controlling nature of the self-view and the reported urge to check if one is depicted correctly to make a good self-presentation. Plots that do not show clear patterns may either contain multiple types of transition groups outlined above or belong to individuals that reportedly purely reflexively looking at themselves. Here, we see the potential for future work to link such patterns to user behavior and compare them across different types of meetings.

5. Differences in self-view perception of cognitive load and control are partially reflected in visual attention and user characteristics

Based on the user's perceptions of cognitive load and control, we identified three user groups that also descriptively differ in their visual behavior (Clusters "Benefiting Users" (C1), "Cognitively Challenged Users" (C2), "Control Losing Users" (C3)). Cluster C2 seems to suffer from the highest cognitive load without benefiting from higher control. The individual's visual behavior shows a descriptively increased median fixation duration and fixation counts compared to C1. Cluster C3 shows a medium level of cognitive load but suffers from low control. These individuals spent the most time interacting with the self-view. Cluster C1 shows only benefits from the self-view (least cognitive load and highest perceived control). These individuals surprisingly interact less with the self-view than others. Increased interaction, especially median fixation duration with the self-view thus may be a driver or proxy for negative aspects of the self-view as a source of cognitive load (C2) or as an element inducing slight cognitive load but not showing positive effects on control (C3). Since the self-view generally provides beneficial effects compared to no self-view in our study, the self-view should be visible for users who use it appropriately but too much interaction seems harmful.

Moreover, our clusters show descriptive differences in terms of neuroticism with C3 individuals having the highest tendency to get nervous or stressed (i.e., high levels of neuroticism). As neuroticism is associated with emotional instability, this result may suggest that these individuals are the ones actively seeking for reassurance by looking at themselves, similar to self-presentation strategies that depend on personality in social network sites [64]. To do so, they may not necessarily often look at themselves, but when they do, spend more time on their self-view compared to others. Beyond, our clusters show descriptive differences in terms of social interaction anxiety with C3 individuals having the highest rating of social interaction anxiety, partially similar to findings from Vriends et al. [90] who showed that highly socially anxious participants did focus often on the self. As visible in the results, individuals clustered in this area do not necessarily look frequently but most prolonged in our study. Combining the findings, we argue that social anxiety-related attention is especially pronounced in tasks that require a strong (negative) focus on self, and lower in a task that is more content-oriented, such as the one we chose. This again underlines the need for a more user and task-conscious self-view design. Lastly, our clusters show descriptive differences in terms of self-satisfaction with C3 individuals having a descriptively lower level of self-satisfaction for "Control Losing Users" (C3). Comparable studies investigating relationships between personality (esp. self-esteem, narcissism, and self-disgust) and gaze show that participants with high self-esteem seem to only briefly look at self-faces and looking at self can decrease self-satisfaction with appearance [7, 75]. In line with their findings, we see descriptive differences that suggest a negative relationship between median fixation duration spent on the self-view and self-satisfaction, however not statistically significant. Since their studies have been conducted in the absence of real social interaction, the descriptive nature may be explained by the fact that individuals can no longer focus on their looks as a main thought. As studies show decreased satisfaction with appearance and an increase in beauty operations assumedly due to a high use of VMS [71, 73], we argue that the effect of high interaction with self-view on self-satisfaction should nevertheless not be underestimated.

8.2 Design Implications

Our study concludes that although the self-view feature is not inherently bad, it cannot be deemed beneficial for everyone. Therefore, the assumption to always turn on the self-view, as done in contemporary VMS, excludes individual user characteristics. We argue that the self-view needs to be designed more flexibly. Therefore, we propose the following design implications:

(1) *Make manual personalization options more prominent:* With manual personalization options, the user's awareness of this feature could be increased by providing an introductory alert message demonstrating this configuration option. As an example, this could be accomplished by making the option to show the self-view a choice in the pre-meeting screen. However, this comes with the cost that individuals need to be aware of the impact the self-view may have on themselves.

(2) *Create an adaptive self-view:* A more advanced way that does not require user interaction could be to design a system-driven adaptive self-view. This design choice symbolizes a so-called user-adaptive system (see e.g., [27]) that may leverage biosignals [80] or further contextual information such as participants familiarity or meeting content. Therefore, a series of possible design approaches can be outlined based on the comments from the interviews and insights of our study.

(a) *Adaptation based on eye-metrics and user groups:* Looking at the combined data of perception and eye-tracking, we see three groups of users varying in their perceived control and cognitive load. These groups also differ in their viewing behavior, especially regarding the median duration spent per fixation on the self-view. Based on these user groups, we propose the following changes: When individuals belong to C2 or C3, we suggest adapting the self-view. This may be instantiated via continuous eye-tracking information and checking whether the fixation duration exceeds the median fixation duration reported for C1 in the results. For those individuals indicating high cognitive load but also high control (C2), we argue that the self-view should be designed less visually stimulating to reduce effects on cognitive load. As these individuals still achieve control, they may be harmed by an eliminated self-view. A default hidden and easily hoverable self-view may be beneficial. For individuals indicating high pre-dispositions for low self-satisfaction, high social anxiety, and neuroticism and showing a visibly lower level of perceived control when intensively interacting with the self-view (C3), we propose eliminating the self-view. These users do not benefit in form of perceived control and show slight increases in cognitive load. One alternative design to ensure higher control for those users could cover image-recognition-based messages that provide reassuring messages. For individuals in C1, we see a beneficial effect on control and a low cognitive load impact. Their self-view may remain static.

(b) *Adaptation based on further contextual information:* Based on the interview insights and related work outlined in Section 2.2, contextual factors of meetings may serve as alternative mechanisms. One possibility could be investigating a meeting participants-dependent self-view based on familiarity with others. Furthermore, we propose a *task-based* self-view. Here, the outlined differentiation between focus and less-focus phases in tasks and focus vs. less-focus tasks could be of interest. Here, the identified gaze data could be used to only show the self-view when participants are looking at the self-view and seem to need it. Thus, stimuli in the peripheral vision that need to be processed and can cause distraction or cognitive load may be reduced in focused phases (see e.g. similar approaches for driver displays in vehicles [68]). A further idea stemming from the interview findings includes the information on how much one is part of the conservation topic, as some interviewees indicated they would like to see themselves while presenting, others don't.

(3) *Change self-view visualization:* Next to dynamically displaying the self-view, its visualization may be changed. For example, applied different filters initially developed to increase user privacy [19] could be used to investigate their effect on the distracting nature of the self-view. This is in line with our identified proposition to apply filtering techniques to the pictures. However, it is important to explore the design space of filters carefully, as especially beauty filters may harm self-satisfaction with appearances despite being possible to apply to the self-view already [72].

8.3 Limitations and Future Work

Although we aimed to follow a rigorous approach, multiple limitations apply.

Experimental Setup. We highlight that our results are only valid in the specific setting and can not be generalized to other situations. Collecting data in a laboratory setting was reasonable as it allowed us to control for confounding factors from the external environment. We acknowledge that our study only presents results for a student sample that was recruited due to their experience with video meetings from online lectures. Therefore, we believe that our sample is also representative of office workers who engage in video meetings due to remote work settings. However, the sample is not comparable in terms of age distribution and potential influencing factors on self-view perception based on age, such as different technology use habits or appearance concerns. We recommend future research to replicate our study with a different task (e.g., a presentation task without a shared screen), and a different sample population (e.g., age, experience with VMS). Lastly, it is important to note that the sample size in our study was limited. Based on the collected eye-tracking data, we followed the sample sizes used in comparable studies that collected neuro- and psychophysiological signals in teams which often have comparably small sample sizes. Therefore, it is important to consider that we can only detect effects of medium size or larger based on our power calculations (one-tailed sensitivity test, alpha error of 0.05, power of 0.8, effect size of d = 0.683). As a result, the significance of effects should be interpreted in light of this effect size.

Self-view and Video Meeting Design. In our study, we decided that one VMS user shares their screen and moderates the content in each round. However, this may induce a bias because there is a difference if a person was a moderator in the first or last round. Additionally, we only include established static self-view designs so far. It might be interesting to extend the study in future research by focusing on other self-view designs and meeting roles, such as a speaker's view.

Observed Gaze Patterns. Despite performing a thorough analysis of low-level eye-tracking metrics, several limitations regarding metrics used, use of multimodal signals, and task variety exist. In the future, further metrics can be calculated. For example, analyzing the similarity of scan paths between participants, or using advanced measurements such as recurrence quantification analysis could provide additional information on inter-personal differences. In addition, combining gaze data and audio recordings could identify the situations in which users interact with the self-view in more detail and provide insights for more task-dependant self-view design opportunities. Comparisons between first and consecutive meetings could better highlight the self-view's effect based on team composition and its impact on self-view's effect, especially in light of impression formation processes within the team and the difference between perceived control in the first round and further rounds for the setting with no self-view. Finally, as gaze patterns are task-dependent [23], it is crucial to understand that our findings only hold for the investigated shared screen and related layout of AOIs. Further studies could compare our identified gaze patterns to gaze patterns in meetings where no shared screen exists (see e.g., George et al. [33]).

Clustering. Finally, it is crucial to note that clustering is an explorative method. To make an informed decision, we applied domain knowledge to check if the clusters made sense and tested their quality with different metrics to select the best clustering. However, the quality of our clustering is limited. We compared the clustering based on gaze data with clusterings based on perception data. Both approaches as well as combining eye-tracking and perception data to cluster lead to the identification of three clusters which are at least partially overlapping (entanglements> 0.5).

9 Conclusion

We investigate the impact of the self-view feature in VMS on the users' cognitive load, perceived control, and the visual behavior in three consecutive meetings of three-person teams with a shared screen. Our study contributes a detailed analysis of the users' perceptions of cognitive load and control as well as their visual behavior patterns in this context. Our findings suggest that users experience higher control when viewing themselves. At the same time, they have a consistently higher cognitive load. Most strikingly, we identified a notable difference in the visual behavior when engaging with the self-view. These differences in visual behavior are also visible in the identified three user groups varying in terms of perception of the cognitive load and perceived control (Benefiting Users, Cognitively Challenged Users, Control Losing Users). The groups also differ in several user characteristics. Especially cluster 2 and 3, describing users that experience disadvantages in form of increased cognitive load or low perceived control despite interacting with the self-view, highlight the dual nature of the self-view. Contrasting these user groups to cluster 1, covering individuals that mostly benefit from the self-view, the varying perceptions of the self-view can be replicated and related to gaze data. By this, we outline detailed recommendations for design improvements. Based on this, we call for designing VMS intelligently and adaptive to the human user to address these different preferences and to reduce the negative effects of the self-view while maintaining its positive impact on control.

Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)-GRK27391-Project Nr. 447089431-Research Training Group: KD School-Designing Adaptive Systems for Economic Decisions. We thank the anonymous reviewers and AC for their valuable input that helped us to improve our manuscript.

References

[1] Olga Abramova, Margarita Gladkaya, and Hanna Krasnova. 2021. An Unusual Encounter with Oneself: Exploring the Impact of Self-view on Online Meeting Outcomes. *International Conference on Information Systems (ICIS) 2021* (2021). https://www.researchgate.net/publication/358395619

[2] Hosam Al-Samarraie. 2019. A Scoping Review of Videoconferencing Systems in Higher Education. *The International Review of Research in Open and Distributed Learning* 20, 3 (feb 2019). https://doi.org/10.19173/IRRODL.V20I4.4037

[3] Hugo J.E.M. Alberts, Carolien Martijn, and Nanne K. de Vries. 2011. Fighting self-control failure: Overcoming ego depletion by increasing self-awareness. *Journal of Experimental Social Psychology* 47, 1 (jan 2011), 58–62. https://doi.org/10.1016/j.jesp.2010.08.004

[4] Jeremy N. Bailenson. 2021. Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior* 2, 1 (feb 2021). https://doi.org/10.1037/tmb0000030

[5] Karolína Balogová and Duncan P Brumby. 2022. How Do You Zoom?: A Survey Study of How Users Configure Video-Conference Tools for Online Meetings. *2022 Symposium on Human-Computer Interaction for Work* (2022). https://doi.org/10.1145/3533406

[6] Boris B. Baltes, Marcus W. Dickson, Michael P. Sherman, Cara C. Bauer, and Jacqueline S. LaGanke. 2002. Computer-Mediated Communication and Group Decision Making: A Meta-Analysis. *undefined* 87, 1 (2002), 156–179. https://doi.org/10.1006/OBHD.2001.2961

[7] Ellise M Barnier and James Collison. 2019. Experimental induction of self-focused attention via mirror gazing: Effects on body image, appraisals, body-focused shame, and self-esteem. *Body Image* 30 (2019), 150–158.

[8] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706QP063OA

[9] Susan A. Brown, Alan R Dennis, and Viswanath Venkatesh. 2010. Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research. *Journal of Management Information Systems* 27, 2 (oct 2010), 9–54. https://doi.org/10.2753/MIS0742-1222270201

[10] Melanie S. Brucks and Jonathan Levav. 2022. Virtual communication curbs creative idea generation. *Nature* 605, 7908 (may 2022), 108–112. https://doi.org/10.1038/s41586-022-04643-y

[11] Roser Cañigueral and Antonia F. de C. Hamilton. 2019. The Role of Eye Gaze During Natural Social Interactions in Typical and Autistic People. *Frontiers in Psychology* 10 (mar 2019). https://doi.org/10.3389/fpsyg.2019.00560

[12] Charles S. Carver and Michael F. Scheier. 1981. Self-consciousness and reactance. *Journal of Research in Personality* 15, 1 (mar 1981), 16–29. https://doi.org/10.1016/0092-6566(81)90003-9

[13] Shelly Chaiken, Akiva Liberman, and Alice H Eagly. 1989. Heuristic and systematic information processing within and beyond the persuasion context. In *Unintended thought.* The Guilford Press, New York, NY, US, 212–252.

[14] Anya Chakraborty and Bhismadev Chakrabarti. 2018. Looking at My Own Face: Visual Processing Strategies in Self–Other Face Recognition. *Frontiers in Psychology* 9 (feb 2018). https://doi.org/10.3389/fpsyg.2018.00121

[15] Shiwei Cheng, Zhiqiang Sun, Xiaojuan Ma, Jodi L. Forlizzi, Scott E. Hudson, and Anind Dey. 2015. Social Eye Tracking: Gaze Recall with Online Crowds. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 454–463. https://doi.org/10.1145/2675133.2675249

[16] Debbie Chew and Mahsa Azizi. 2022. The State of Video Conferencing 2022 | Dialpad. https://www.dialpad.com/blog/video-conferencing-report/

[17] Marvin M. Chun, Julie D. Golomb, and Nicholas B. Turk-Browne. 2011. A Taxonomy of External and Internal Attention. *Annual Review of Psychology* 62, 1 (jan 2011), 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

[18] Edwin S. Dalmaijer, Sebastiaan Mathôt, and Stefan Van der Stigchel. 2014. PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behavior Research Methods* 46, 4 (dec 2014), 913–921. https://doi.org/10.3758/s13428-013-0422-2

[19] Jose Eurico De Vasconcelos Filho, Kori M. Inkpen, and Mary Czerwinski. 2009. Image, appearance and vanity in the use of media spaces and videoconference systems. *GROUP'09 - Proceedings of the 2009 ACM SIGCHI International Conference on Supporting Group Work* (2009), 253–261. https://doi.org/10.1145/1531674.1531712

[20] Rachel del Valle. 2020. Why Are All Video Call Platforms Designed with the "Self-view?" – Eye on Design. https://eyeondesign.aiga.org/why-are-all-video-chat-platforms-designed-with-the-self-view/

[21] Angelika Dimoka, Rajiv D. Banker, Izak Benbasat, Fred D. Davis, Alan R. Dennis, David Gefen, Alok Gupta, Anja Ischebeck, Peter H. Kenning, Paul A. Pavlou, Gernot Müller-Putz, René Riedl, Jan Vom Brocke, and Bernd Weber. 2012. On the use of neurophysiological tools in is research: Developing a research agenda for neurois. *MIS Quarterly: Management Information Systems* 36, 3 (2012), 679–702. https://doi.org/10.2307/41703475

[22] Nicola Döring, Katrien De Moor, Markus Fiedler, Katrin Schoenenberg, and Alexander Raake. 2022. Videoconference Fatigue: A Conceptual Analysis. *International Journal of Environmental Research and Public Health* 19, 4 (feb 2022), 2061 – 2081. https://doi.org/10.3390/ijerph19042061

[23] Andrew T. Duchowski. 2007. *Eye Tracking Methodology: Theory and Practice* (2 ed.). Springer London, London. 334 pages. https://doi.org/10.1007/978-1-84628-609-4
[24] S Duval and R.A. Wicklund. 1972. *A theory of objective self awareness.* Academic Press, New York. https://psycnet.apa.org/record/1973-26817-000
[25] Carmen Egido. 1988. Video conferencing as a technology to support group work: a review of its failures. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work - CSCW '88*. ACM Press, New York, New York, USA, 13–24. https://doi.org/10.1145/62266.62268
[26] Thompson et al. 2004. NASA Exercise: Survival on the Moon. (2004).
[27] Stephen Fairclough and Kiel Gilleade. 2012. Construction of the biocybernetic loop. In *Proceedings of the 14th ACM international conference on Multimodal interaction - ICMI '12*. ACM Press, New York, New York, USA, 571. https://doi.org/10.1145/2388676.2388797
[28] Geraldine Fauville, Mufan Luo, Anna Carolina Muller Queiroz, Jeremy N. Bailenson, and Jeff Hancock. 2021. Nonverbal Mechanisms Predict Zoom Fatigue and Explain Why Women Experience Higher Levels than Men. *SSRN Electronic Journal* (apr 2021). https://doi.org/10.2139/SSRN.3820035
[29] Allan Fenigstein. 1987. On the Nature of Public and Private Self-Consciousness. *Journal of Personality* 55, 3 (sep 1987), 543–554. https://doi.org/10.1111/J.1467-6494.1987.TB00450.X
[30] Mark Fenton-O'Creevy, Nigel Nicholson, Emma Soane, and Paul Willman. 2003. Trading on illusions: Unrealistic perceptions of control and trading performance. *Journal of Occupational and Organizational Psychology* 76 (3 2003), 53–68. Issue 1.
[31] A. Field, J. Miles, and Z. Field. 2012. *Discovering Statistics Using R.* SAGE Publications. https://books.google.de/books?id=wd2K2zC3swIC
[32] Liz Fosslien and Mollie West Duffy. 2020. How to Combat Zoom Fatigue. https://hbr.org/2020/04/how-to-combat-zoom-fatigue
[33] Joey George, Akmal Mirsadikov, Misty Nabors, and Kent Marett. 2022. What do Users Actually Look at During 'Zoom' Meetings? Discovery Research on Attention, Gender and Distraction Effects. In *Proceedings of the Hawaii International Conference on System Sciences*. Manoa, Hawaii. https://doi.org/10.24251/HICSS.2022.582
[34] John M. Govern and Lisa A. Marsch. 2001. Development and Validation of the Situational Self-Awareness Scale. *Consciousness and Cognition* 10, 3 (sep 2001), 366–378. https://doi.org/10.1006/CCOG.2001.0506
[35] Jerald Greenberg. 1983. Self-image versus impressional management in adherence to distributive justice standards: The influence of self-awareness and self-consciousness. *Journal of Personality and Social Psychology* 44, 1 (jan 1983), 5–19. https://doi.org/10.1037/0022-3514.44.1.5
[36] Jay Hall and W. H. Watson. 1970. The Effects of a Normative Intervention on Group Decision-Making Performance. *Human Relations* 23, 4 (aug 1970), 299–317. https://doi.org/10.1177/001872677002300404
[37] Qiaohong Hao, Mateu Sbert, and Lijing Ma. 2019. Gaze Information Channel in Cognitive Comprehension of Poster Reading. *Entropy* 21, 5 (apr 2019), 444. https://doi.org/10.3390/e21050444
[38] SG Hart and LE Staveland. 1988. Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In *Human mental workload*, PA Hancock and N Meshkati (Eds.). Amsterdam, Netherlands, 139–83.
[39] Martin Hassell and John Cotton. 2017. Some things are better left unseen: Toward more effective communication and team performance in video-mediated interactions. *Computers in Human Behavior* 73 (aug 2017), 200–208. https://doi.org/10.1016/J.CHB.2017.03.039
[40] Andrea B Hollingshead and Joseph E McGrath. 1995. Computer-assisted groups: A critical review of the empirical research. *Team effectiveness and decision making in organizations* (1995), 46–78.
[41] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer. 2011. *Eye Tracking: A comprehensive guide to methods and measures.* OUP Oxford. https://books.google.de/books?id=5rIDPV1EoLUC
[42] Ryan Horn and Tara Behrend. 2017. Video killed the interview star: Does picture-in-picture affect interview performance? *Personnel Assessment and Decisions* 3, 1 (may 2017), 5. https://doi.org/10.25035/pad.2017.005
[43] Gernot Horstmann and Arvid Herwig. 2016. Novelty biases attention and gaze in a surprise trial. *Attention, Perception, & Psychophysics* 78, 1 (jan 2016), 69–77. https://doi.org/10.3758/s13414-015-0995-1
[44] Daniel R. Ilgen, John R. Hollenbeck, Michael Johnson, and Dustin Jundt. 2005. Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology* 56 (2005), 517–543. https://doi.org/10.1146/ANNUREV.PSYCH.56.091103.070250
[45] Mansoor Iqbal. 2022. Zoom Revenue and Usage Statistics 2022. https://www.businessofapps.com/data/zoom-statistics/
[46] Robert J. K. Jacob and Keith S. Karn. 2003. Eye Tracking in Human-Computer Interaction and Usability Research.
[47] R Johansen. 1988. *Groupware: Computer Support for Business Teams.* Free Press, New York, NY.
[48] Aditi Jublie and Devpriya Kumar. 2021. Early Capture of Attention by Self-Face: Investigation Using a Temporal Order Judgment Task. *i-Perception* 12, 4 (jul 2021), 204166952110329. https://doi.org/10.1177/20416695211032993

[49] Nancy Kanwisher and Galit Yovel. 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences* 361, 1476 (dec 2006), 2109–2128. https://doi.org/10.1098/rstb.2006.1934

[50] Katherine A. Karl, Joy V. Peluchette, and Navid Aghakhani. 2022. Virtual Work Meetings During the COVID-19 Pandemic: The Good, Bad, and Ugly. *Small Group Research* 53, 3 (jun 2022), 343–365. https://doi.org/10.1177/10464964211015286

[51] Alboukadel Kassambara. 2017. *Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning (Multivariate Analysis) - Softcover.* CreateSpace Independent Publishing Platform.

[52] Andy J. King, Nadine Bol, R. Glenn Cummins, and Kevin K. John. 2019. Improving Visual Behavior Research in Communication Science: An Overview, Review, and Reporting Recommendations for Using Eye-Tracking Methods. *Communication Methods and Measures* 13, 3 (jul 2019), 149–177. https://doi.org/10.1080/19312458.2018.1558194

[53] René F. Kizilcec, Kathryn Papadopoulos, and Lalida Sritanyaratana. 2014. Showing face in video instruction: Effects on information retention, visual attention, and affect. *Conference on Human Factors in Computing Systems - Proceedings* (2014), 2095–2102. https://doi.org/10.1145/2556288.2557207

[54] Ned Kock. 2004. The psychobiological model: Towards a new theory of computer-mediated communication based on Darwinian evolution. *Organization Science* 15, 3 (2004), 327–348. https://doi.org/10.1287/orsc.1040.0071

[55] Ned Kock. 2011. Media naturalness theory: human evolution and behaviour towards electronic communication technologies. *Applied evolutionary psychology* (2011), 381–398.

[56] Daniel G Krutka and Kenneth T Carano. 2016. Videoconferencing for Global Citizenship Education: Wise Practices for Social Studies Educators. *Journal of Social Studies Education Research* 7, 2 (2016), 109–136. www.sosyalbilgiler.org

[57] Kristine M. Kuhn. 2022. The constant mirror. *Computers in Human Behavior* 128 (mar 2022), 107110. https://doi.org/10.1016/J.CHB.2021.107110

[58] Anastasia Kuzminykh and Sean Rintel. 2020. Classification of Functional Attention in Video Meetings. *Conference on Human Factors in Computing Systems - Proceedings* (apr 2020). https://doi.org/10.1145/3313831.3376546

[59] J Lafferty. 1974. *The Desert survival problem : a group decision making experience for examining and increasing individual and team effectiveness: manual.* Experimental Learning Methods, Plymouth Mich.

[60] Moritz Langner, Peyman Toreini, and Alexander Maedche. 2022. EyeMeet: A Joint Attention Support System for Remote Meetings. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts.* ACM, New York, NY, USA, 1–7. https://doi.org/10.1145/3491101.3519792

[61] Mark R. Leary and Robin M. Kowalski. 1990. Impression management: A literature review and two-component model. *Psychological Bulletin* 107, 1 (jan 1990), 34–47. https://doi.org/10.1037/0033-2909.107.1.34

[62] Yuan Hang Li and Nim Tottenham. 2011. Seeing yourself helps you see others. *Emotion* 11, 5 (2011), 1235–1241. https://doi.org/10.1037/a0023469

[63] Thomas Maran, Marco Furtner, Simon Liegl, Theo Ravet-Brown, Lucas Haraped, and Pierre Sachse. 2021. Visual Attention in Real-World Conversation: Gaze Patterns Are Modulated by Communication and Group Size. *Applied Psychology* 70, 4 (oct 2021), 1602–1627. https://doi.org/10.1111/apps.12291

[64] Minas Michikyan, Kaveri Subrahmanyam, and Jessica Dennis. 2014. Can you tell who I am? Neuroticism, extraversion, and online self-presentation among young adults. *Computers in Human Behavior* 33 (apr 2014), 179–183. https://doi.org/10.1016/j.chb.2014.01.010

[65] Matthew K. Miller and Martin Dechant. 2021. Meeting you, seeing me: The role of social anxiety, visual feedback, and interface layout in a get-to-know-you task via video chat. *Conference on Human Factors in Computing Systems - Proceedings* (may 2021). https://doi.org/10.1145/3411764.3445664

[66] Matthew K Miller, Regan L Mandryk, Max V Birk, Ansgar E Depping, and Tushita Patel. 2017. Through the Looking Glass: Effects of Feedback on Self-Awareness and Conversation during Video Chat. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017). https://doi.org/10.1145/3025453

[67] Akmal Mirsadikov and Joey George. 2023. Can you see me lying? Investigating the role of deception on gaze behavior. *International Journal of Human-Computer Studies* 174 (2023), 103010. https://doi.org/10.1016/j.ijhcs.2023.103010

[68] T Miura. 1986. Coping with situational demands: A study of eye movements and peripheral vision performance. *Vision in vehicles* (1986).

[69] Annabel Ngien and Bernie Hogan. 2022. The relationship between Zoom use with the camera on and Zoom fatigue: considering self-monitoring and social interaction anxiety. *https://doi.org/10.1080/1369118X.2022.2065214* (2022). https://doi.org/10.1080/1369118X.2022.2065214

[70] Jakob Nielsen. 2006. Ten Usability Heuristics.

[71] Gabrielle N. Pfund, Jennifer Harriger, and Patrick L. Hill. 2021. Video chat usage and the Big Five in women during the COVID-19 pandemic. *Personality and Individual Differences* 171 (mar 2021), 110537. https://doi.org/10.1016/J.PAID.2020.110537

[72] Gabrielle N. Pfund, Patrick L. Hill, and Jennifer Harriger. 2020. Video chatting and appearance satisfaction during COVID-19: Appearance comparisons and self-objectification as moderators. *International Journal of Eating Disorders* 53, 12 (dec 2020), 2038–2043. https://doi.org/10.1002/EAT.23393

[73] Toni D Pikoos, Simone Buzwell, Gemma Sharp, and Susan L Rossell. 2021. The Zoom Effect: Exploring the Impact of Video Calling on Appearance Dissatisfaction and Interest in Aesthetic Treatment During the COVID-19 Pandemic. *Aesthetic Surgery Journal* 41, 12 (nov 2021), NP2066–NP2075. https://doi.org/10.1093/asj/sjab257

[74] Michael I. Posner. 1980. Orienting of Attention. *Quarterly Journal of Experimental Psychology* 32, 1 (feb 1980), 3–25. https://doi.org/10.1080/00335558008248231

[75] Jonas Potthoff and Anne Schienle. 2021. Effects of Self-Esteem on Self-Viewing: An Eye-Tracking Investigation on Mirror Gazing. *Behavioral Sciences* 11, 164 (nov 2021), 1–10. https://doi.org/10.3390/bs11120164

[76] Kemper C J Klein M C Beierlein C & Kovaleva A Rammstedt B. 2014. Big Five Inventory (BFI-10). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)* (2014). https://doi.org/10.6102/zis76

[77] Falko Rheinberg, Regina Vollmeyer, and Stefan Engeser. 2003. Die erfassung des flow-erlebens. (2003).

[78] René Riedl. 2021. On the stress potential of videoconferencing: definition and root causes of Zoom fatigue. *Electronic Markets 2021* 1 (dec 2021), 1–25. https://doi.org/10.1007/S12525-021-00501-3

[79] Bertrand Schneider and Roy Pea. 2013. Real-time mutual gaze perception enhances collaborative learning and collaboration quality. *International Journal of Computer-Supported Collaborative Learning* 8, 4 (dec 2013), 375–397. https://doi.org/10.1007/S11412-013-9181-4/TABLES/1

[80] Tanja Schultz and Alexander Maedche. 2023. Biosignals meet Adaptive Systems. *SN Applied Sciences* 5, 9 (09 Aug 2023), 234. https://doi.org/10.1007/s42452-023-05412-w

[81] Julia Seitz, Ivo Benke, Armin Heinzl, and Alexander Maedche. 2024. The Impact of Video Meeting Systems on Psychological User States: a State-of-the-Art Review. *International Journal of Human-Computer Studies* 182 (2024), 103178. https://doi.org/10.1016/j.ijhcs.2023.103178

[82] Kristen M. Shockley, Allison S. Gabriel, Daron Robertson, Christopher C. Rosen, Nitya Chawla, Mahira L. Ganster, and Maira E. Ezerins. 2021. The fatiguing effects of camera use in virtual meetings: A within-person field experiment. *Journal of Applied Psychology* 106, 8 (aug 2021), 1137–1155. https://doi.org/10.1037/apl0000948

[83] Abdulla Al Suman, Carlo Russo, Ann Carrigan, Patrick Nalepka, Benoit Liquet-Weiland, Robert Ahadizad Newport, Poonam Kumari, and Antonio Di Ieva. 2021. Spatial and time domain analysis of eye-tracking data during screening of brain magnetic resonance images. *PLOS ONE* 16, 12 (dec 2021), e0260717. https://doi.org/10.1371/journal.pone.0260717

[84] Zoom Support. 2023. See or Hide My Video. https://support.zoom.com/hc/en/article?id=zm_kb&sysparm_article=KB0068249

[85] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285.

[86] Peyman Toreini, Moritz Langner, and Alexander Maedche. 2020. Using Eye-Tracking for Visual Attention Feedback. In *Lecture Notes in Information Systems and Organisation*. Vol. 32. Springer, 261–270. https://doi.org/10.1007/978-3-030-28144-1_29

[87] Ralf A. L. F. van Griethuijsen, Michiel W. van Eijck, Helen Haste, Perry J. den Brok, Nigel C. Skinner, Nasser Mansour, Ayse Savran Gencer, and Saouma BouJaoude. 2015. Global Patterns in Students' Views of Science and Interest in Science. *Research in Science Education* 45, 4 (aug 2015), 581–603. https://doi.org/10.1007/s11165-014-9438-6

[88] Margot van Wermeskerken and Tamara van Gog. 2017. Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning. *Computers & Education* 113 (oct 2017), 98–107. https://doi.org/10.1016/j.compedu.2017.05.013

[89] Kathleen D. Vohs, Roy F. Baumeister, and Natalie J. Ciarocco. 2005. Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology* 88, 4 (apr 2005), 632–657. https://doi.org/10.1037/0022-3514.88.4.632

[90] Noortje Vriends, Yasemin Meral, Javier A. Bargas-Avila, Christina Stadler, and Susan M. Bögels. 2017. How do I look? Self-focused attention during a video chat of women with social anxiety (disorder). *Behaviour Research and Therapy* 92 (may 2017), 77–86. https://doi.org/10.1016/J.BRAT.2017.02.008

[91] Hana Vrzakova, Mary Jean Amon, McKenzie Rees, Myrthe Faber, and Sidney D'Mello. 2021. Looking for a Deal? *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (jan 2021), 1–35. https://doi.org/10.1145/3434169

[92] Hana Vrzakova, Mary Jean Amon, Angela E B Stewart, and Sidney K D'Mello. 2019. *Dynamics of Visual Attention in Multiparty Collaborative Problem Solving Using Multidimensional Recurrence Quantification Analysis*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300572

[93] Jürgen Wegge. 2006. Communication via videoconference: Emotional and cognitive consequences of affective personality dispositions, seeing one's own picture, and disturbing events. *Human-Computer Interaction* 21, 3 (2006), 273–318. https://doi.org/10.1207/S15327051HCI2103_1

[94] Steve Whittaker. 1995. Rethinking video as a technology for interpersonal communications: theory and design implications. *International Journal of Human-Computer Studies* 42, 5 (may 1995), 501–529. https://doi.org/10.1006/

IJHC.1995.1022

[95] Han Zhang, Kevin F. Miller, Xin Sun, and Kai S. Cortina. 2020. Wandering eyes: Eye movements during mind wandering in video lectures. *Applied Cognitive Psychology* 34, 2 (mar 2020), 449–464. https://doi.org/10.1002/ACP.3632

[96] Zoom. 2022. App Marketplace. https://marketplace.zoom.us/

A Appendix

Table 5. Self-created or modified scales used per survey

Survey	Construct	Code	Question	Scale	Factor Load
Pre, DV	Cognitive Load	CL001	How mentally demanded do you feel at this moment?	10 item likert	0.75, 0.42, 0.47, 0.56
		CL002	How physically demanded do you feel at this moment?		0.92, 0.67, 0.74, 0.95
		CL003	How rushed or hurried do you feel at this moment?		0.33, 0.64, 0.63, 0.62
		CL004	How inscecure, discouraged, irritated, stressed or annoyed do you feel at this moment?		-, 0.71, 0.88, 0.49
DV	Control	PC001	The video meeting software allowed me to control the information I sent to the others about myself (e.g. myself, my appearance and behavior).	7 item likert	0.34, 0.49, 0.58
		PC002	I felt powerless to control the information I sent to the others about myself (e.g. myself, my appearance and behavior).		0.73, 0.78, 0.70
		PC003	In the video meeting, the information I sent to the others about myself (e.g. myself, my appearance and behavior) was completely out of my hands.		0.82, 0.73, 0.77
		PC004	My information sent to others on my behavior, appearance and myself keeps the same during the video meeting, regardless of what I do.		-0.26, -0.34, -0.54
	Selfview Frequency	PV001	How often did you look at your own image?	7 item likert	0.61, 0.64, 0.45
Post	Gender	Gender	Please indicate your gender:	f/m/d	1 (fixed)
	Age	Age	Please indicate your age:	range	1 (fixed)
	General Condition	GPC001	My eyes are physically in a good condition	yes/ no	1 (fixed)
		GPC003	Did you wear visual aids during the experiment? (e.g. contact lenses, glasses)	yes/ no	1 (fixed)
	Self-satisfaction with appearance	SS001	How satisfied are you with your face today?	5 item likert	0.46
		SS002	How satisfied are you with your face in general?		0.66
		SS003	How satisfied are you with your body today?		0.68
		SS004	How satisfied are you with your body in general?		0.79
	Experience Task	ET001	I am experienced in the topic of astronautic	7 item likert	1 (fixed)
		ET002	I am experienced in the topic of sailing		1 (fixed)
		ET003	I am experienced in the topic of camping/ surviving in unknown terrain		1 (fixed)
	Experience Meeting	EM001	I am experienced with videoconferences and virtual meetings.	7 item likert	1 (fixed)
	SCS-social anxiety		as in original, SCS005, SCS009, SCS011, SCS017	7 item likert	0.77, 0.31, 0.41, 0.37
	SCS-public		as in original, SCS004, SCS007, SCS015, SCS020	7 item likert	0.46, 0.65, 0.79, 0.76
	SCS-private		as in original, SCS006, SCS012	7 item likert	0.63, 0.41
	BFI -extra		as in original, BFI001, BFI006	7 item likert	0.03, 0.05
	BFI -neuro		as in original, BFI004, BFI009	7 item likert	0.35, 0.45

Table 6. Overview on cronbach alpha values per construct

Construct	Cronbach alpha
Cognitive Load 0	0.836
Cognitive Load 1	0.711
Cognitive Load 2	0.64
Cognitive Load 3	0.705
Control 1	0.219
Control 2	0.503
Control 3	0.29
Control modified (shortened) 1	0.699
Control modified (shortened) 2	0.802
Control modified (shortened) 3	0.803
Experience with meeting	1 (only 1 item)
Experience with task	1 (only 1 item)
Age	1 (only 1 item)
Gender	1 (only 1 item)
Self-Consciousness-Scale private	0.422
Self-Consciousness-Scale public	0.643
Self-Consciousness-Scale situation anxiety	0.551
Self satisfaction (with appearance)	0.745
Personality (BFI Neuro)	0.583
Personality (BFI Extraversion)	0.817

Received July 2023; revised January 2024; accepted March 2024

Table 7. Overview of themes and codes - User states

Theme	Codes of theme	Explanation	Example quote
Control	Level of control (High/Low)	Level of Control coded as high or low	High: "I did have the feeling that I have a certain amount of control, simply because I'm in front of the camera and I can always check whether everything fits" (S1:a28) Low: "it would have been a little bit helpful to see oneself. and then see how that comes over when the hands are quiet, you know" (S2:a16 (S2:a16)
	Control mechanism	Mechanism used to achieve control coded as via self-view or others	Self-view: "When I saw myself on the screen, I had the opportunity to look again, how is your hair sitting there somehow, do you look a little strange or so" (S1:a21) Others: "of course, you always have control over what you say and stuff, the image that the other person had of you is always very subjective as well" (S1:a9)
	Control coping (coping, no coping)	Coping mechanism whether having no self-view as control is applied or not	Coping: "I didn't really have control over the picture, but I was a little bit careful to sit up straight, not slouching on the chair or moving back and forth" (S2:a15) No coping: "it would have been a little bit helpful to see oneself" (S2:a16) "because I did not see myself and only saw the others, I lost myself as an image, so to speak, that I was now sitting there quite normally (...)" (S2:a34).
Cognitive load	Level of Cognitive Load	Level of Cognitive Load coded as high or low	High: no quotes available Low: "were not mentally demanding tasks or not a challenging one " (S1:a36)
Temporal changes	Changes of Control	Changes in control between meetings, coded as stable or increasing or decreasing	Stable: "So basically I would say a constant sense of control" (S1:a Increasing: no quotes available Decreasing: "I made sure that the camera wasn't standing in any inappropriate or unflattering way in general. But at the beginning, it was just a matter of ten seconds. And after that, I didn't really pay any attention to it" (S2:a16)
	Changes of Cognitive Load	Changes in cognitive load between meetings, coded as stable or increasing or decreasing	Stable: "the tasks (were) all very doable, I would say. And... Well, I wasn't particularly strained there" (S2:a16) Increasing: "in the third meeting it was just, I think, because of the length of the whole session the most challenging meeting but just because of the tiredness that came up then" (S2:a22) Decreasing: "and then on the second and third run it was used to it and that's why it was then also, such a low load"(S2:a15) "somehow I felt like when I shared my screen (cognitive load) was the highest, (...) but I don't think it's because I shared it, but because it was an exercise that we had more thoughts about" (S1:a18

Table 8. Overview of themes and codes - Self-view interaction

Theme	Code of theme	Explanation	Example quote
Temporal changes	Changes of Self-view Behavior	Changes in self-view behavior between meetings, coded as stable or changing	Stable, Constant no pattern: made an effort to not look at myself "as soon as you had the time or the opportunity to look at, you tend to do that" (S1:a21) Changing - most when focused: "you do not know the participants and you don't want them to go out there and think, oh, she was handful or she (...) was totally stupid. That is why I thought about it. I think it became less, (...) then I didn't have the doubts anymore" (S1:a14) Changing - most when no focus: "didn't feel like I paid that much attention during the productive meetings" (S1:a14, S1:a31) "mainly during waiting periods" (S1:a14) "It was more when maybe a few things were still missing or also when we were done and discussed a bit afterwards. But so during, so now when it came to remembering this season, I wasn't really paying attention to me then" (s1:a9)
	Changes of Visual attention	Changes in attention and focus between meetings, coded as increase or decrease in a AOI	Increase in people and decrease on content: you discussed and looked again with the others, is there something else that you want to change? And you still tended to look at the others more, even if the other person didn't notice it that way" (S1:a21)
Self-attention	Self-attention focus area (Participants, Self-view, Content	Self-attention and concentration focus based on described area of interest	Content: "my focus was still on the content and also on what the others said" (S1:a25) "completely dissociated from my own appearance" (S2:a32) "good (to not see oneself), because you pretty much sink into the task. I actually found that really pleasant" (S2:a19) People: "I was more concentrated on other people's decisions, so I was you know looking on the others images screen when interacting in the task" (s1:a36) Self: "as soon as you had the time or the opportunity to look at, (...) because, yes, I think, you tend to do that. It is somehow a reflex when you see yourself that you look" (S1:a21)
	Degree of Self-view Behavior	Degree of self-view behavior between meetings, coded as no, limited, intense	No-self/ Low: "made an effort to not look at myself" (S1:a31) Limited focus on self: if (you) start focusing more on the meeting, (you) will just forget" (S1:a31) Intense focus on self: "as soon as you had the time or the opportunity to look at, you tend to do that" (S1:a21) "definitely looked at (themselves)" (S1:a28)

Table 9. Overview of themes and codes - Self-view design

Theme	Code of theme	Explanation	Example quote
Self-view in general	Perception of self-view (positive, negative, neutral)	Perception of having a self-view coded as positive, negative or neutral	Positive: as soon as you had the time or the opportunity to look at, (...) because, yes, I think, you tend to do that. It is somehow a reflex when you see each other that you look" (S1:a21) Negative: "made an effort to not look at myself" (S1:a31) Neutral: "it makes it neither better or worse" (S1:a25)
	Perception of no self-view (Positive, Negative, Neutral)	Perception of having no self-view coded as positive, negative or neutral	Positive: "good (to not see oneself), because you pretty much sink into the task. I actually found that really pleasant" (S2:a19) "no self-view, (...) increased the immersion into the group effort and the attention" (S2:a2) Negative: "there was no picture of me, that made me a bit insecure, because I'm sure I'm not in a camera, whether my face is all the way in and so on" (S2:a38) Neutral: "uncommon to not see my image in the software, but that was no problem" (S2:a2)
Self-view Design Ideas	Alternative Design (positive/negative)	Attitude towards alternative design visualizations coded based on positive or negative	Positive: " Yes, I could imagine that. Maybe it makes more sense than a static one because you're not as constantly focused on yourself." (S1:a9) " I think it's a very exciting idea to take away self-view when I'm not talking or should be focused. Because you're simply not the center of the group dynamic and then you don't need any feedback and thoughts about how you appear." (S2:a22) Negative: "So in general I don't think it would bother me, I just don't know what the advantage would be, so if you have one more function to see yourself, then I don't know why you should opt for software that has less functionality but is otherwise just as good" (S2:a15)
	Alternative Design mechanism	Proposed alternative design visualizations coded based on their name	Adaptable: "could also imagine that there is also a separate picture or there could be two other switch buttons or something like (hiding the self-view)" (S1:a14) Eye-based: "That means maybe a pop-up, where you could see your self-view or something like that based on the gaze data, so that if you want to look, you can look, but not by chance so the gaze wanders to oneself." (S2:a11) Task-based: "if it's a PowerPoint presentation or something, then it would be good to have an alternate or hidden view" (S1:a9) Participant based: "more interested when being with friends in how I present myself" (S2:a19)
	Alternative Design visualization	Proposed alternative design visualizations coded based on their name	Filter: "apply compression methods to make the picture more fluid" (S2:a32)

Table 10. Summary statistics of eye-tracking metrics on self-view for design variant 1

Construct	Statistical Value	Round 1	Round 2	Round 3
Share of fixation counts on AOI (%)	Min	00.118	00.224	0
	Q1	00.530	00.389	00.574
	Median	00.8609	01.439	00.811
	Mean	02.681	02.718	02.494
	Q3	02.730	04.730	02.771
	Max	16.667	09.453	16.101
Share of time spent on AOI (%)	Min	00.041	00.122	0
	Q1	00.3582	00.297	0.310
	Median	00.871	01.274	01.309
	Mean	03.185	02.534	02.557
	Q3	03.314	04.529	02.674
	Max	22.0241	12.096	17.595
Count of fixations (number)	Min	3	5	0
	Q1	13	8	11
	Median	19	33.5	27
	Mean	62.294	61.9444	57.588
	Q3	62	103	59
	Max	374	214	384
Min fixation duration (ms)	Min	55.515	55.515	0
	Q1	55.515	55.514	55.515
	Median	55.515	55.515	55.515
	Mean	57.475	57.983	59.434
	Q3	55.516	55.515	55.5163
	Max	66.619	77.721	99.928
Maximum fixation duration (ms)	Min	66.619	177.650	0
	Q1	299.784	421.919	399.712
	Median	899.353	738.358	821.631
	Mean	956.828	755.013	929.397
	Q3	1232.446	977.075	1110.312
	Max	2753.575	1487.819	2686.956
Mean fixation duration (ms)	Min	66.618	122.134	0
	Q1	134.0313	154.056	163.771
	Median	237.924	196.276	198.052
	Mean	220.125	203.318	200.436
	Q3	265.879	246.119	233.166
	Max	463.247	312.754	413.428
Median fixation duration (ms)	Min	66.618	99.928	0
	Q1	116.583	122.135	127.686
	Median	144.340	141.564	160.995
	Mean	146.947	153.902	150.5453
	Q3	172.098	177.650	166.547
	Max	233.0610	255.372	321.989