Lack of Eye Contact Detection Thresholds–A Two-Alternative Forced Choice Study

BreAzia Echols(✉) and Voicu Popescu

Purdue University, West Lafayette, IN 47906, USA
{echolsb,popescu}@purdue.edu

Abstract. The ability to establish and sustain eye contact is a prerequisite for effective communication. Unfortunately, eye contact is difficult to preserve in remote collaboration that relies on videoconferencing, mixed reality, or virtual reality technologies. Before proceeding with the design of remote collaboration systems that provide adequate eye contact, an important preliminary question is how good does eye contact have to be for a remote collaborator to not feel that their interlocutor is not making eye contact. This paper presents a two-alternative forced choice (2AFC) study with N = 139 participants that measured the lateral detection threshold for lack of eye contact. The findings suggest that a person sitting in front of a laptop has to have a fixation point within 1.5 cm (or 1.43°) of the laptop's video camera for their remote interlocutor to not perceive the lack of eye contact.

Keywords: Computing Methodologies · Computer Graphics · Graphics Systems and Interfaces · Mixed/Augmented Reality

1 Introduction

Advances in technology have made real-time remote collaboration ubiquitous. The appearance and voice of each user is acquired, transmitted, and rendered in real time, abstracting away geographic distances. Cameras have sufficient resolution, networks have sufficient bandwidth, and computers have sufficient processing power to enable effective real time collaboration across continents. Such remote collaboration is now used routinely in many domains, including education [7,21], healthcare [14,29], and social connectedness [4,10].

Remote collaboration solutions differ through how participants are acquired [8]. One option is to acquire participants in 3D, which has the advantages of allowing the viewer to perceive their remote collaborator with accurate depth cues, of integrating the remote participant seamlessly into the viewer's physical surroundings or into a collaborative virtual environment with fine grain depth compositing, and of allowing the viewer to change the viewpoint from where they see their remote collaborator [26]. The main challenge of the approach is that it is difficult to acquire a person in 3D in real time and with sufficient color and geometry resolution to allow for a realistic rendering at the remote sites [25]. Another option is to acquire the participants in 2D, with a

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
D. Michael-Grigoriou et al. (Eds.): EuroXR 2025, LNCS 16101, pp. 317–330, 2026.
https://doi.org/10.1007/978-3-032-03805-0_17

conventional video camera. The advantages are acquisition simplicity–most laptops, tablets, and phones do have a user-facing videocamera–and visualization realism. The challenges of remote collaboration based on 2D acquisition are the lack of depth cues, the difficulty of seamless integration into 3D physical or virtual environments, the limited range of viewpoints from where the 2D acquired participant can be seen, and the difficulty of supporting eye contact between a participant and their remote interlocutor.

Most remote collaboration is conducted through videoconferencing. Videoconferencing is the simplest form of 2D remote collaboration, where each participant sees a real time video of each of the remote participants. Videoconferencing ignores all challenges enumerated above. A participant does not see the remote participants in 3D, the remote participants are not integrated into the 3D physical space surrounding the local participant or in a collaborative 3D virtual environment, and participants do not make eye contact with one another. Whereas the shortcomings stemming from the lack of 3D data are palatable for many remote collaboration applications, the lack of eye contact is a fundamental limitation of videoconferencing that greatly reduces the effectiveness of remote collaboration [24]. Without eye contact, a participant is likely to find it harder to follow what their interlocutor says, the conversation turns to a series of monologues as opposed to a dialogue, and participants disengage [2].



Fig. 1. Frames from pairs of videos like the ones shown to participants in our study. For the first pair (top), the offsets for *1a* and *1b* are 0 cm and 7 cm, respectively. For the second pair (bottom), the offsets for *2a* and *2b* are 1 cm and 0 cm, respectively.

Consider the scenario when participant A and participant B engage in videoconferencing, each using their laptop. If A looks on their laptop at the eyes of B,

then A looks away from their laptop's webcam, and therefore A's video does not provide eye contact to B. For B to feel that A is looking at them, A would have to look at their webcam, which, of course, precludes A from seeing B. While there have been attempts to remedy the lack of eye contact in videoconferencing, prior work has not answered reliably the important first question of how good we are at detecting the lack of eye contact. In other words, what is the threshold above which eye contact in videoconferencing is detected by participants? Estimating this threshold would provide a target to any approach aimed at addressing the issue of eye contact in videoconferencing and in remote collaboration in general.

In this paper we present a study (N = 139) that measured the detection threshold for lack of eye contact in video conferencing. The study was approved by our institution's ethics board, i.e., the Institutional Review Board (IRB). The study follows a two alternative forced choice (2AFC) design, which is the golden standard for psycho-physical detection threshold studies. A participant is shown pairs of videos of a potential remote interlocutor, one at the time, and is asked which of the two videos had better eye contact. In one of the videos the interlocutor looks straight at their camera, and in the other the interlocutor looks away from their camera, by various offsets. Figure 1 shows frames from two pairs of videos. Whereas it is straightforward to choose *1a* over *1b* as the frame with better eye contact, choosing between *2a* and *2b* is harder.

2 Prior Work

2.1 The Lack of Eye Contact Problem

The problem of lack of "mutual gaze" in video conferencing has been noted for a long time [28]. The recent pandemic that disrupted all aspects of our lives and had us heavily rely on video conferencing triggered an intense research interest in video conference effectiveness and lack thereof [9]. Although eye contact is used sparingly in normal conversations, being able to establish eye contact when needed is essential for communication. For example, eye contact has been shown to influence problem solving and to facilitate conceptual understanding [18]. Full gaze awareness is the ability to know at what someone is looking, and is an important component when considering the necessities of utilizing videoconferencing platforms. Mutual gaze is the occurrence of two people making direct eye contact with each other. It is often difficult to achieve both full gaze awareness and mutual gaze in videoconferencing environments, due to the vertical discrepancy between the camera and the facial image that causes offset eye contact to be made [22].

2.2 Attempts to Address the Eye Contact Problem

Many solutions have been proposed to alleviate the problem of lack of eye contact in video conferencing, but none have found widespread use. Videoconferencing proceeds by and large without trying to address the problem. One solution is to rely on the user of the videoconferencing system to learn when their remote

interlocutors are looking at them [12]. The user calibrates the various focus points and remembers which view direction the remote interlocutor assumes when they are looking at the user. The result is of course just the most indirect of eye contacts, and what is learned becomes obsolete when the videoconferencing participants–or the videoconferencing system itself–rearrange where each participant is displayed. Another simple solution is to rotate the plane on which the video is displayed to place the head of the viewer on the view direction axis of the person in the video [17]. The shortcoming is that the method lacks the benefit of the 3D geometry of the person's head, so the video becomes the video of a video projected on a slanted plane, and not the video of someone talking into the camera.

Neural networks rival humans in detecting eye contact [6], but neural network gaze manipulation is still primarily an offline process, unsuitable for real-time remote collaboration [11]. Neural networks have indeed achieved real time performance only by limiting the gaze redirection to the eyeballs, i.e., excluding the face [15]. However, real-time 3D gaze redirection continues to rely on modeling and re-rendering [13]. The work aimed at addressing the lack of eye contact is orthogonal to ours. Our focus is to measure the lack of eye contact detection thresholds, which informs any and all approaches for addressing the lack of eye contact.

Beyond videoconferencing, eye contact has long been, and continues to be, an important design concern in collaborative mixed reality [1,3]. Compared to videoconferencing, collaborative mixed reality has the advantage of greater flexibility in placing the visualization of the remote collaborator, as well as the benefit of a richer representation of the remote collaborator, which often includes 3D data. One study examined the effectiveness of a remote collaborative system where some participants are seated at a traditional computer, whereas another participant was in a mixed reality environment where the other collaborators were virtually placed around a table. One of the main benefits to the participant in the mixed reality environment was the ability to better perceive nonverbal cues, including natural gaze cues [3]. This highlights the importance of conveying eye contact correctly in collaborative XR systems, and our study gathers important preliminary data informing the design of such systems. Our work, conducted in the context of videoconferencing, does inform all remote collaboration approaches regarding eye contact requirements.

2.3 Lack of Eye Contact Detectability

Prior research on eye contact sensitivity is limited. One study examined four setups for video tele-consultation to find, unsurprisingly, that a 7° eye gaze angle yielded better perceived eye contact than a 14° angle [27]. Some studies relied on rendered stimuli based on virtual 3D head models, and others on video stimuli. A dense exploration of the gaze directions space was conducted with the help of a virtual 3D head, which was aimed away from the study participant with increasing angular offsets [20]; the study found that the cone of gaze directions from which a person feels looked at subtends 8.3° in the lateral direction. A

subsequent study, also relying on a virtual 3D head, found that the region of perceived eye contact is symmetrical in all directions, except for a slight 0.5° asymmetry downwards, and that it is 6.8° wide [23].

Sensitivity to eye contact was measured with video stimuli in a study that used eight directions, i.e., every 45° [5], and fixation offsets between 0° and 5° in all directions except for downwards, which was between 0° and 15°, with 1° increments. More than half of the participants perceived eye contact when the offset was approximately 1°, except for the downward direction, where the offset was 7°. This substantial asymmetry, which was not replicated by the later study of Van der Pol [23], prompted a second replication study that found that the sensitivity to eye contact is 1.2° horizontally and 1.7° vertically.

Another study was performed where the recorded subject focused their gaze on different offsets, and respondents were asked for a series of still-frame screenshots if the subject was looking directly into the camera. This study found that in order for respondents to achieve 75% accuracy, an offset of around 2.7° laterally was required. [19]

2.4 Anchoring Our Work Into Prior Work

Building on this prior work, our study contributes novel insights on eye contact sensitivity. First, prior work [5,19,20,23] employs an experimental design that relies on yes-no questions, e.g., "Do you feel looked at, yes or no?" or "Do you feel like the person is making eye contact, yes or no?". Such a design, sometimes called a pseudo 2AFC design, is more prone to bias than a true 2AFC design. For example, instead of answering the question whether the person in the video is making eye contact, participants might answer the question "Do you feel like the eye contact is good enough?", which depends on individual preference. On the other hand, in a true 2AFC design a participant will not choose the offset option over the no offset option just because they find it acceptable. In another example, a participant in a light detection study might truthfully answer "no" when presented with a dim light stimulus, although if that same stimulus were presented paired with a stimulus with no light at all, the participant might still be able to choose the dimmer stimulus reliably.

Similarly, in a pseudo 2AFC design, the order in which the videos are presented has a much larger impact, as the participant will likely end up rating later videos that they are presented with relative to the videos that they had already seen up until that point. In other words, in a pseudo 2AFC design, the participant's standards can change throughout the experiment due to the subjectivity of the question, whereas in a true 2AFC study such as ours, the user will always only be considering whether each offset video has more or less direct eye contact than the video it is paired with (where the video subject is looking directly at the camera). This leads to far less variance in each users' responses, resulting in more accurate results compared to prior works.

Our study employs a 2AFC design, which is the best design for threshold detection studies. Eye contact detection is a quick task that allows many repetitions per participant as needed to investigate a wide range of offsets, so a 2AFC design is possible and should be used, unlike, for example, in the case of gain detection threshold studies for redirected walking in virtual reality.

A second way in which our study is complementary to prior work is that we investigate lack of eye contact detection thresholds not in a laboratory setting, with stimuli recorded with a camera at the height of the user and with fixation points on the vertical plane [5,20,23], but rather in the realistic setting of a remote collaborator seated in front of their laptop, with the laptop screen angled as desired, and with fixation points around the laptop's webcam, which provides application validity.

3 Study

In order to fulfill the need for a detectability threshold at which individuals are able to notice a lack of eye contact, we performed an experiment to measure response rates in order to calculate such a threshold. Our study was approved by our IRB. The IRB protocol number will be included in the non-anonymous version of this paper.

3.1 Methods

Study Design. Our study follows a two-alternative forced choice (2AFC) design. As detailed in Sect. 2, the 2AFC design is the best known design for detection threshold measurement, as it less prone to bias compared to the yes-no question (i.e., "pseudo 2AFC") design employed in prior work [5,19,20,23].

We show participants pairs of short videos, each three seconds in duration. For each pair, in one video the person in the video is looking straight at the camera, and in the other video they are looking at a point that is offset with respect to the camera (Fig. 1). All video pairs have one video with no offset, and one with an offset in the 0 cm–10 cm range. The two videos of a pair are shown one at the time, in random order. After the participant has seen both videos of a pair, they are asked which video had better eye contact: the first or the second. The detection threshold is measured as the offset for which the average correct response rate is 75%, i.e., half way between chance (50%) and perfect (100%).

The focus of this study was limited to examining the detectability threshold of horizontal offsets. Part of the justification for this was to limit participant fatigue, as just the addition of examining vertical offsets would double the number of trials that each participant would be asked to undergo. Increased participant fatigue would likely lead to less accurate results and potentially a lower rate of users completing the entire study. An alternative method would be to divide the participants into subgroups which are each asked to undergo trials on only one direction, but this would incur further shortcomings, such as reduced number of participants per direction, and the possibility of each subgroup being imbalanced

in terms of eye dominance or vision ability. Focusing on fewer directions also allows for greater granularity and range in the offsets presented to users, whereas prior studies presented far fewer videos to participants in the horizontal direction than our study, such as [5] which for the horizontal direction only operated on a range of 0° to 5°, with a granularity of 1° between videos, whereas our study had approximately double the granularity and range, leading to far more accurate results.

Focusing on the detectability of the horizontal thresholds is also logical since, during video conferencing, users are more likely to have windows of other interlocutors located on the left or right sides of their screens (meaning that their own gaze will be offset as they focus on the window of the other interlocutor). Furthermore, prior research has indicated that humans are more sensitive to detecting horizontal offsets than they are to detecting vertical offsets.

Participants. Our study had $N = 139$ participants from our university's community. Based on the demographics data we have collected, 37.4% of participants were female, 61.9% were male, and 0.72% checked the "prefer not to disclose" option. Furthermore, 84.2% of participants were in the 18–25 age range, 14.4% were in the 26–35 range, and 1.44% were in the over 35 range. Finally, 12.2% of participants indicated that they use videoconferencing daily, 48.2% use it at least once a week, 26.6% use it at least once a month, 12.2% use it at least once a year, and 0.72% had never used it. Participants were not screened for color vision deficiency, but the subject of the video has dark irises, so color vision deficiency should be less impactful than otherwise.

Stimuli. We recorded videos in a typical video conferencing setting, with the same person seated in front of their laptop. Each video was 3 s long and the person was asked to say the same line: "Hi, my name is <name>". The videos differed through the person's fixation point, i.e., the point where they looked when the video was recorded. We recorded 21 videos, one with a fixation point at the laptop's webcam, i.e., an offset of 0 cm, the next with a fixation point 0.5 cm to the right, the next with a fixation point of 1.0 cm to the right of the laptop's webcam, and so on, with the last video having a fixation point of 10.0 cm to the right of the laptop webcam (Fig. 2). The 21 videos were then mirrored left-right in software to create 21 left offset videos. Note that the 0 cm offset video was also mirrored to avoid lighting differences with the left offset images. The procedure created 42 videos that were shown to participants according to the following procedure.

Mirroring the videos from a single direction was done for consistency, as any minor asymmetries in the video subject's face, lighting, dominant eye, or environment could impact the viewer's perception of eye contact. These variables should be kept constant since we are only concerned with how the change in offset affect's the viewer's perception of eye contact, to reduce bias.

Also for consistency, the videos were all recorded by a single individual (an author of the study, to ensure accuracy of offset). There are many variables

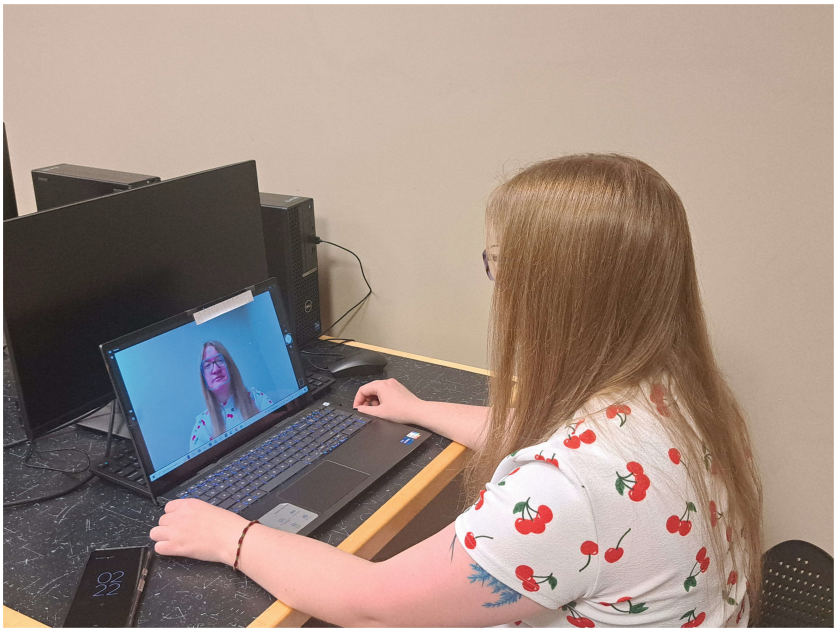


Fig. 2. Back lateral image of the setup used for the recording of the video stimuli.

of an individual's appearance (ethnicity, race, gender, age, glasses, etc.) which may affect a participant's ability to compare eye contact between two videos, and varying any of these attributes between videos would lead to inconsistencies in the users' selections, and attempting to cover enough variations of these attributes so as to not weaken the study would require an intractable number of participants to adjust for the variability between which variation they were presented with.

Since this variable could not be removed, it should be noted that the subject of the video wears glasses, as do 63.7% of individuals in the USA and over 4 billion people across the globe, meaning that it is meaningfully representative for the subject to have glasses.

Experimental Procedure. The study was run online, through Qualtrics. First a participant was prompted for demographics data and then the participant was shown pairs of videos and asked to choose the video with the better eye contact. The study took approximately 10–15 min per participant. Each participant was compensated with a cash gift card worth the equivalent of $5.

Each participant was presented with 42 pairs of videos, covering all possible left and right offsets. The left offset pairs were shown together, as were the right offset pairs. Compared to alternating left and right pairs, our approach of showing videos of the same type of offset in block translates to less variability between consecutive pairs. This in turn helps participants detect lack of eye con-

tact, and it results in more conservative detection thresholds. Some participants saw the left-offset videos first, and some last, in counterbalanced fashion. For each type of offset, left or right, the 21 pairs covered all 21 offsets, from 0 cm to 10 cm, with 0.5 cm increment, with one of the two videos being the 0 cm video. The offsets were shown out of order, in counterbalanced fashion. After each pair, the participant was asked to identify the video with the better eye contact. The participant had to provide an answer to move on to the next pair.

Data Collection. In addition to the demographics data, we recorded for each participant and for each pair of videos shown to them whether the participant identified correctly the video with better eye contact. This amounts to 42 bits of data per participant. For the two video pairs, one left and one right, where both videos were the 0 cm offset video, the accuracy was defined as being 50%, since the two stimuli are perceptually indifferentiable, so the respondent is choosing at random between the two videos. This definition is consistent with 2AFC studies, where the point at which the stimuli can not be differentiated is defined to have an accuracy of 50%.

Data Analysis. We analyzed the data in three steps.

In a first step we computed the correct answer rate for each offset as a percentage. The correct answer rate for an offset is defined as the number participants who answered correctly for that offset over the total number of participants. This first step yields 21 data points for the left offsets, and 21 for the right offsets. Each 2D data point is defined by an offset and the correct answer rate at that offset.

In a second step we fit a sigmoid psychometric function to the 21 left offset data points, and one to the 21 right offset data points, using a least squares optimization. The sigmoid function, whose use is standard practice in 2AFC experiments, is given in Eq. 1.

$$S(x, \alpha, \beta, \gamma, \lambda) = 0.5 + (0.5 - \lambda)\left(\frac{1}{1 + e^{\gamma \frac{x-\alpha}{\beta}}}\right) \tag{1}$$

Here, x is the function's variable, i.e., the offset, and α, β, γ and λ are parameters set by the fitting process. We use γ to distinguish between the left and the right offset sigmoids. We choose the left offsets to be negative, i.e., from 0 cm to -10 cm, and the right offsets to be positive, i.e., from 0 cm to 10 cm. Therefore the left sigmoid decreases as the offset increases in algebraic sense from -10 cm to 0 cm, and $\gamma = 1$. The right sigmoid increases as the offset increases from 0 cm to 10 cm, and $\gamma = -1$. α, β and λ are set through the least squares optimization. α controls the lateral shift, β controls the lateral scale, and γ controls the maximum value, as data is unlikely to have a perfect response rate of 100%, even for large offsets.

In a third step, each of the two sigmoids was intersected with the y = 75% line to estimate the left offset and right offset detection thresholds.

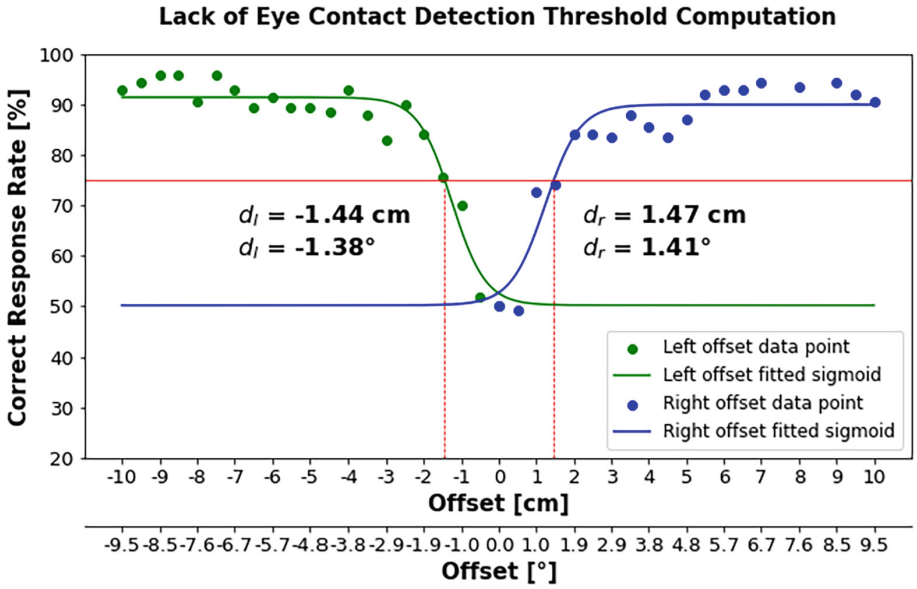


Fig. 3. Experimental data (dots), psychometric sigmoid functions fitted to the experimental data (curves), and detection thresholds (d_l, d_r), for both left (green) and right (blue) eye contact offsets. A data point (x, y) gives the correct response rate y at offset x. The detection thresholds are computed by intersecting each sigmoid with the $y = 75\%$ line. (Color figure online)

3.2 Results and Discussion

The experimental data is shown in Fig. 3. For small offsets, the correct response rate is close to chance behavior, i.e., 50%. For large offsets, participants identify the video with the better contact reliably, i.e., with correct response rates over 90%. Even for the largest offsets, the correct response rate is not 100%, which we explain as participants forgetting, or remembering incorrectly, which video had no eye contact offset.

The least squares optimization yields the following parameters (see Eq. 1) for the left offset sigmoid: $\lambda_l = 0.0863$, $\alpha_l = 1.25$, $\beta_l = 0.449$. The detection threshold d_l is 1.44 cm, or 1.38°. The fitting error is 0.0348, or 3.48%, which yields a detection threshold confidence interval of [1.441cm, 1.447cm] or [1.376°, 1.381°].

For the right offset sigmoid, the fitted parameter values are: $\lambda_r = 0.101$, $\alpha_r = 1.24$, $\beta_r = 0.458$. The detection threshold d_r is 1.47 cm, or 1.40°. The fitting error is 0.0447, or 4.47%, which yields a detection threshold confidence interval of [1.468 cm, 1.477 cm] or [1.401°, 1.411°].

These offset thresholds above which participants can reliably detect lack of eye contact are small. This is in line with some prior work, which also find detection thresholds in the 1° range [5], though it differs from other prior work which found a lateral detection threshold of 2.7° [19], for which we provide a

potential explanation of difference in methodology (pseudo 2AFC for that study as opposed to true 2AFC for ours), as well as sample size ($N = 56$ for that study as opposed to $N = 139$ for ours). The small detection thresholds imply demanding targets for procedures aimed at remedying the lack of eye contact.

We have also estimated detection thresholds for the various sub-populations of our study. The results are shown in Table 1. The detection thresholds do not change much with gender or age. The only sub-population for which the thresholds are substantially different from those of the overall population is the sub-population of participants who use videoconferencing daily. For these participants the detection thresholds are much smaller, which is plausible since frequent users of videoconferencing technology might also have grown better at telling when their remote interlocutor is making eye contact, and when they are not. However, there are only 17 participants in this category, which is a relatively small number. Furthermore, the fitting error is also larger than for other categories, so additional experiments have to be conducted to confirm whether frequent users of videoconferencing technology do indeed have lower detection thresholds for lack of eye contact.

Table 1. Fitting error ε and detection threshold d in cm and in degrees, for both left (subscript l) and right (subscript r) offsets, for the various sub-populations of our study.

Sub-population	ε_l[%]	ε_r [%]	d_l [cm]	d_r [cm]	d_l [°]	d_r [°]
Female	4.12	5.75	−1.56	1.85	−1.49	1.77
Male	3.96	5.56	−1.34	1.42	−1.28	1.35
Age 18–25	3.47	4.52	−1.45	1.45	−1.38	1.39
Age >25	6.63	6.99	−1.39	1.55	−1.33	1.48
Daily	7.89	8.59	−0.83	0.53	−0.79	0.51
Weekly	3.27	5.80	−1.51	1.01	−1.44	0.96
Monthly	6.03	4.99	−1.64	1.86	−1.56	1.78
Yearly/Never	5.11	6.99	−1.30	1.41	−1.25	1.34

4 Conclusion, Limitations and Future Work

We have conducted a study for measuring lack of eye contact detection thresholds in a typical videoconferencing scenario with a remote participant seated in front of their laptop and being acquired by the laptop's videocamera. We have used a 2AFC design which is robust to bias in participant answers. We have collected data from a large number of participants (N = 139), which has resulted in a robust estimation of the detection thresholds, with small fitting errors. The results show that the user has to look within approximately 1.5 cm of their

laptop videocamera for their collaborator to not detect the lack of eye contact. At 60 cm this translates to a stringent 1.5°. Our work informs the design of remote collaboration systems, whether based on videoconferencing or on other approaches, such as mixed or virtual reality, to make video sprites or avatars of remote collaborators convey eye contact adequately.

One limitation of our study is that most data points are for offsets that are well above the detection threshold. Indeed, for offsets with an absolute value of 3 or above, the detection rate is already quite large. A finer determination of the detection thresholds would require more data points in the 0 cm to 3 cm range. This is difficult to achieve with the current experimental apparatus, as it is difficult to record videos focusing on fixation points on a scale with 0.1 cm increments. Future work could look into implementing the fixation point as a dot moving on the laptop screen, which it can do with pixel or sub-pixel accuracy, providing a finer grain sampling of the offset range. As always, participant fatigue, which can manifest itself through mechanical answers, is a concern that competes with the potential benefits brought by more trials with finer offset differences.

Another limitation of our study is that we focused exclusively on left-right offsets, to minimize participant fatigue. Most laptops have their integrated webcam placed at the top of their screen, so users do not look above the camera when they look at the screen. Future studies could investigate detection thresholds for a downward offset, although, as prior work as shown, we are much better at detecting lateral offsets than downward offsets. This means that the lateral offset detection thresholds found in our study are a conservative estimate of the detection thresholds for downward offsets.

Finally, our study was performed using only a single actor in the presented video. While this was done for the sake of consistency between videos, this does lead to the limitation that this could impede generalizability, as certain physical characteristics (e.g. gender, race, eye color and size, facial structure, etc.) may influence a viewer's perception of eye contact. Ideally, more studies would be done in the future to verify that our measured thresholds are consistent across multiple demographics.

The lack of eye contact detection thresholds will be especially useful in informing the design of extended reality systems. Consider, for example, the scenario of remote lecture attendance, where the instructor wears an XR headset to see remote students rendered into the empty seats of a classroom [16]. Each remote student is seated in front of their laptop and is acquired by the laptop's webcam. The thresholds will inform on the quality of the eye contact perceived by the instructor and also on the amount the remote student video sprites have to be rotated to sustain eye contact, based on the classroom empty seat that accommodates each remote student.

It is plausible that our society's reliance on remote collaboration will only continue to increase. In order to satisfy this demand, and to make remote collaboration more effective, future work could propose and test computer screen designs that extend the display function with an acquisition function. A sim-

ple such display could integrate camera islands between the display pixels, and videoconferencing software could place each remote collaborator's video at one such camera island. This way, when collaborator A looks on screen at remote collaborator B, A will also direct their gaze towards the camera that acquires A for B, conveying to B that A is making eye contact with them. One challenge is that each participant will be acquired multiple times, once for each of their collaborators, which leads to a proliferation of video streams that have to be transmitted. As camera miniaturization continues to advance, the islands of "dead pixels" will continue to shrink, with the acquisition function not coming at the cost of a diminished display function.

Acknowledgements. The authors wish to thank Yuqi Zhou and Sarah Engle for their help, as well as our numerous user study participants without whom this study would not have been possible. This material is based upon work supported by the United States National Science Foundation under Awards No. 2318657, 2309564, and 2212200.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bai, H., Sasikumar, P., Yang, J., Billinghurst, M.: A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–13. CHI '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376550
2. Basch, J.M., Melchers, K.G., Kurz, A., Krieger, M., Miller, L.: It takes more than a good camera: which factors contribute to differences between face-to-face interviews and videoconference interviews regarding performance ratings and interviewee perceptions? J. Bus. Psychol. **36**, 921–940 (2021)
3. Billinghurst, M., Kato, H.: Collaborative mixed reality. In: Proceedings of the First International Symposium on Mixed Reality, pp. 261–284 (1999)
4. Bleakley, A., et al.: Bridging social distance during social distancing: exploring social talk and remote collegiality in video conferencing. Hum. Comput. Interact. **37**(5), 404–432 (2022)
5. Chen, M.: Leveraging the asymmetric sensitivity of eye contact for videoconference. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 49–56 (2002)
6. Chong, E., et al.: Detection of eye contact with deep neural networks is as accurate as human experts. Nat. Commun. **11**(1), 6386 (2020)
7. Correia, A.P., Liu, C., Xu, F.: Evaluating videoconferencing systems for the quality of the educational experience. Distance Educ. **41**(4), 429–452 (2020)
8. Druta, R., Druta, C., Negirla, P., Silea, I.: A review on methods and systems for remote collaboration. Appl. Sci. **11**(21), 10035 (2021)
9. Fauville, G., Luo, M., Queiroz, A., Lee, A., Bailenson, J., Hancock, J.: Videoconferencing usage dynamics and nonverbal mechanisms exacerbate zoom fatigue, particularly for women. Comput. Hum. Behav. Rep. **10**, 100271 (2023)

10. Gan, Y., Greiffenhagen, C., Reeves, S.: Connecting distributed families: camera work for three-party mobile video calls. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–12 (2020)
11. Ganin, Y., Kononenko, D., Sungatullina, D., Lempitsky, V.: DeepWarp: photorealistic image Resynthesis for gaze manipulation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9906, pp. 311–326. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46475-6_20
12. Grayson, D.M., Monk, A.F.: Are you looking at me? Eye contact and desktop video conferencing. ACM Trans. Comput. Hum. Interact. (TOCHI) **10**(3), 221–243 (2003)
13. Guo, Y., Zhang, J., Chen, Y., Cai, H., Huang, Z., Deng, B.: Real-time face view correction for front-facing cameras. Comput. Visual Media **7**(4), 437–452 (2021). https://doi.org/10.1007/s41095-021-0215-y
14. Haleem, A., Javaid, M., Singh, R.P., Suman, R.: Telemedicine for healthcare: capabilities, features, barriers, and applications. Sens. Int. **2**, 100117 (2021)
15. Hsu, C.F., Wang, Y.S., Lei, C.L., Chen, K.T.: Look at me! Correcting eye gaze in live video communication. ACM Trans. Multimedia Comput. Commun. Appl. (TOMM) **15**(2), 1–21 (2019)
16. Huang, S., Popescu, V.: HyperXRC: hybrid in-person+ remote extended reality classroom-a design study. In: 2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR), pp. 609–618. IEEE (2024)
17. Jaklič, A., Solina, F., Šajn, L.: User interface for a better eye contact in videoconferencing. Displays **46**, 25–36 (2017)
18. Joiner, R., Scanlon, E., O'Shea, T., Smith, R.B., Blake, C.: Evidence from a series of experiments on videomediated collaboration: Does eye contact matter? In: Computer Support for Collaborative Learning, pp. 371–378. Routledge (2023)
19. Kushner, B.J.: Eccentric gaze as a possible cause of "zoom fatigue." J. Binocular Vision Ocular Motility **71**(4), 175–180 (2021)
20. Macrae, C.N., Hood, B.M., Milne, A.B., Rowe, A.C., Mason, M.F.: Are you looking at me? Eye gaze and person perception. Psychol. Sci. **13**(5), 460–464 (2002)
21. Martin, M.: Seeing is believing: the role of videoconferencing in distance learning. Br. J. Edu. Technol. **36**(3), 397–405 (2005)
22. Monk, A.F., Gale, C.: A look is worth a thousand words: full gaze awareness in video-mediated conversation. Discourse Process. **33**(3), 257–278 (2002). https://doi.org/10.1207/S15326950DP3303_4
23. van der Pol, D.: The effect of slant on the perception of eye-contact. Master's thesis, Eindhoven University of Technology (2009)
24. Riedl, R.: On the stress potential of videoconferencing: definition and root causes of zoom fatigue. Electron. Mark. **32**(1), 153–177 (2022)
25. Rokhsaritalemi, S., Sadeghi-Niaraki, A., Choi, S.M.: A review on mixed reality: current trends, challenges and prospects. Appl. Sci. **10**(2), 636 (2020)
26. Schäfer, A., Reis, G., Stricker, D.: A survey on synchronous augmented, virtual, and mixed reality remote collaboration systems. ACM Comput. Surv. **55**(6), 1–27 (2022)
27. Tam, T., Cafazzo, J.A., Seto, E., Salenieks, M.E., Rossos, P.G.: Perception of eye contact in video teleconsultation. J. Telemed. Telecare **13**(1), 35–39 (2007)
28. Vertegaal, R., van der Veer, G.C., Vons, H.: Effects of gaze on multiparty mediated communication (2000). https://api.semanticscholar.org/CorpusID:7393589
29. Viswanathan, R., Myers, M.F., Fanous, A.H.: Support groups and individual mental health care via video conferencing for frontline clinicians during the COVID-19 pandemic. Psychosomatics **61**(5), 538–543 (2020)