Check for updates

Visual attention and cognitive effects of facial anonymization in 360° videos

Leslie Wöhler[1]*, Satoshi Ikehata[2] and Kiyoharu Aizawa[1]

[1]Department of Information and Communication Engineering, The University of Tokyo, Tokyo, Japan, [2]Digital Content and Media Sciences Research Division, National Institute of Informatics, Tokyo, Japan

We analyze perceptual effects of facial anonymization in 360° videos to understand how different anonymization technologies affect the attention and cognition of viewers. 360° videos provide highly immersive viewing experiences, making them ideal for the exploration of real world scenes. As videos of public locations generally include bystanders, it is necessary to apply facial anonymization to protect their privacy. However, this changes the visual content and might affect the attention and cognition of viewers. To investigate these effects, we perform an experiment in which participants watch 360° videos while we collect their eye tracking data. Additionally, we prepare questionnaires measuring presence, video quality, and memory of participants. Our results show differences in visual attention and the perceived video quality between anonymization techniques highlighting that it is important to consider the chosen facial anonymization technique.

KEYWORDS

facial anonymization, face-swapping, visual perception, virtual reality, 360° video, Deepfake

1 Introduction

Immersive viewing experiences allow users to explore scenes with a high sense of realism and presence. When viewing real world videos in head-mounted displays (HMDs), users can feel as if they are actually transported to a different location making the technology highly interesting for tourism and education where it can be used to explore different cities (Takenawa et al., 2023), plan travel (Kumar et al., 2022), and conduct virtual field trips (Thomas Ruberto et al., 2023). Unfortunately, recording in public places inevitably leads to the filming of bystanders which incurs privacy concerns (Faklaris et al., 2020). While the privacy of bystanders can be protected by applying facial anonymization, the technology changes the video content and can affect the perception of viewers (Wöhler et al., 2024; Khamis et al., 2022; Li et al., 2017). Despite investigations on this topic, it is still unclear whether facial anonymization in immersive viewing scenarios affects the attention and cognition of viewers. Therefore, we conduct an experiment that uses eye tracking to measure visual attention and questionnaires to understand the memorization ability of participants watching anonymized and non-anonymized 360° videos.

Our experiment complements previous work that investigates facial anonymization focusing on viewing on regular screens (Li et al., 2017; Khamis et al., 2022; Wilson et al., 2022) as well as augmented reality devices Corbett et al. (2023) and HMDs (Wöhler et al., 2024). In general, studies of regular videos and images on screens identified that

FIGURE 1
An exemplar frame taken from a video recording a dance parade. We study three types of anonymization in 360° videos: Blocking (left), blurring (middle), and replacement of facial areas by face-swapping (right).

anonymization techniques with a higher degree of realism like replacement with cartoon avatars or face-swapping are highly effective and popular with viewers (Hassan et al., 2017; Khamis et al., 2022). Furthermore, even the anonymized bystanders did perceive face-swapping anonymization as highly effective and natural (Khamis et al., 2024). For 360° videos, a previous study suggests differences in the perception of facial anonymization between viewing on regular screen and HMD (Wöhler et al., 2024). The authors state that facial anonymization of 360° videos viewed in HMD can have a negative effect on presence and feel distracting. These results highlight the impact of facial anonymization in immersive viewing scenarios; however, it remains unclear whether facial anonymization actually affects the cognition and attention of viewers. So far, the cognition of immersive content has been studied for various topics finding that users can confuse the source of memories between the real and virtual world (Bonnail et al., 2024) and that the ability to memorize content is affected by the appearance of avatars (Mizuho et al., 2024). Furthermore, learning outcomes can differ between screens and HMD with better results in HMD despite higher cognitive load (Chao et al., 2023). The results can be further improved by guiding the attention of viewers (Liu et al., 2022). Therefore, the literature suggests that various factors can influence presence, memory, and attention in immersive viewing scenarios. This leads to the question of how facial anonymization can affect the visual attention of viewers as well as their ability to memorize video content.

To gain a better understanding of how facial anonymization in immersive environments affect cognitive processes, we measure the eye gaze of viewers watching 360° videos in an HMD. Following a previous study, we use 360° videos of various street scenes filmed in Tokyo which are anonymized by hiding faces behind black boxes, applying gaussian blur to the facial area, and replacing original faces with a synthetic appearance using face-swapping (Wöhler et al., 2024), see Figure 1. In contrast to the previous study, we do not inform the participants about our intention and do not discuss facial anonymization before the experiment. This allows us to obtain natural, non-biased eye tracking data. After each trial, we ask participants to report their presence on the IPQ questionnaire (Regenbrecht and Schubert, 2002; Schubert et al., 2001; Schubert, 2003) as well as their impression of the video quality. After all trials, we show images to participants and ask them to decide whether an image corresponds to content that they saw during the experiment. This way, we aim to measure whether facial anonymization impacts

the ability of participants to remember scene content as this can be highly relevant for videos in domains like marketing or education. Finally, we perform a debriefing with participants following a structured interview which assesses their opinion of the different anonymization techniques.

Based on our results, we answer the following research questions:

- Is the visual attention of participants affected by facial anonymization?
- Is the ability of participants to memorize the videos affected by facial anonymization?
- Is facial anonymization perceived to reduce the video quality?

2 Methods

We assessed our research questions by conducting a perceptual experiment showing various 360° videos with and without facial anonymization to participants on an HMD.

2.1 Hypotheses

As previous work found that blurring and blocking had a negative effect on presence and scene impression of participants in experiments where participants were aware of facial anonymization (Wöhler et al. 2024), we hypothesized that the same effects will be present in our experiment. Following this, we also assumed that especially blocking will lead to a negative impression of the video quality. As people do not necessarily notice face-swapping in videos (Wöhler et al., 2024; Wöhler et al., 2021), we assumed that the viewing behavior as well as memory of participants will be more similar between face-swapped and non-anonymized videos.

In summary, we assessed the following hypotheses:

- H1: Face-swapped videos will convey more presence than blurred or blocked videos. The scene impression of blocked videos will be worse than for the other anonymization techniques.
- H2: Facial anonymization will negatively impact the perceived video quality.
- H3: Gazing behavior is affected by facial anonymization.

- H4: Facial anonymization can reduce the ability to remember the videos.

2.2 Stimuli

The stimuli consisted of 360˚ videos recorded along public roads which have been previously used to study facial anonymization (Wöhler et al., 2024). This allows us to easily relate our findings to previous work. The videos show popular shopping districts, tourist attractions, as well as dance performances in Tokyo. Three types of facial anonymization have been applied to videos: Blocking of facial areas with black rectangles, blurring with gaussian blur, and face-swapping utilizing the SimSwap framework (Chen et al., 2020). The videos do not include audio. An example of the anonymizations is shown in Figure 1.

During the experiment, stimuli were assigned to participants using a counter-balanced between-participant design with randomization. While every participant watched each of the 16 videos, the condition (Non-anonymized, Block, Blur, Face-swap) was chosen at random. Furthermore, we made sure that each participant saw exactly four videos for each condition. This way, we obtained the same number of annotations for each video-condition pair.

To investigate whether facial anonymization can impact the ability to remember scene content, we prepare 12 image stimuli. Six of those are taken from the videos shown to participants, while six show different scenes. The images are cropped from the full panorama taken around the initial viewing position to ensure they were seen by participants.

2.3 Apparatus and measures

The videos were shown to the participants on a PICO 4 Enterprise HMD which is equipped with an eye tracking system (72 Hz). Before the experiment, we adjusted the inter pupillary distance to suit the participant and performed 9-point eye tracking calibration using the provided software of the HMD. During the experiment, participants were asked to focus on a fixation cross before each trial to ensure a consistent initial viewing position. We performed drift correction between trials. While watching the videos, participants were seated on a rotating office chair allowing them to easily turn and watch the whole surrounding of the video.

Next to the eye tracking data, we collected data using questionnaires. Following previous work (Wöhler et al., 2024), participants were answering questions on their scene impression and their sense of presence using the IPQ questionnaire (Regenbrecht and Schubert, 2002; Schubert et al., 2001; Schubert, 2003). Additionally, we added four questions about the video quality to measure whether facial anonymization could be perceived to lower the video quality. We asked participants how they would rate the overall video quality, the sharpness, colors, and whether the video contained artifacts. All questions used a seven point Likert-scale.

After all trials, participants viewed images and decided whether they showed content of any of the videos they saw during the experiment. Additionally, we performed a debriefing with the participants following a structured interview. Here, we first asked participants whether they noticed that facial anonymization was applied to the videos. If they answered yes, we asked which types of anonymization they recognized and their opinions about them. Afterwards, we briefly explained face-swapping and asked whether they noticed that it was used. If participants noticed face-swapping, we asked them about their impression of this technique.

2.4 Participants

We recruited 20 participants for our experiment using an online recruiting site. The participants consisted of eight female and 12 male participants with an average age of 31.15 years (SD 10.35). All participants were Japanese nationals currently living in Tokyo or the neighboring prefectures. All participants had normal or corrected to normal vision using contact lenses. The experiment took around 45 min and participants were rewarded with a 1,500 Yen gift card.

2.5 Procedure

In contrast to previous work (Wöhler et al., 2024), we did not inform the participants about the usage of facial anonymization. This allowed us to collect non-biased eye tracking data and investigate more natural viewing behavior than in previous studies. At the start of the experiment, we instructed participants, obtained informed consent, and collected demographic information. Furthermore, we demonstrated how to use the HMD and performed eye tracking calibration.

During the experiment, participants first viewed a fixation cross for 3 s, afterwards the randomly chosen stimuli video was played automatically. Participants were encouraged to rotate their chair to change their viewing direction. They were not able to pause or replay the video. Once the video finished, a questionnaire was displayed. After answering all questions, the next fixation cross was shown.

After all 16 trials were completed, the post experiment questionnaire was displayed. Here, participants were shown one image at a time and asked whether this image was taken from one of the videos they saw. The experiment ended with a structured interview as debriefing.

3 Results

3.1 Questionnaire data

To analyze the questionnaire data, we first verified that the assumptions of normality and the homogeneity of variances are satisfied using Shapiro-Wilk and Levene tests. Afterwards, we visualized the questionnaire data as bar plots, see Figure 2.

Against our hypothesis H1, we find no significant differences between the anonymization conditions for the IPQ presence questionnaire (ANOVA $F_{(3, 4476)} = 0.97$, $p > 0.4$) or scene impression (ANOVA $F_{(3, 2236)} = 1.14$, $p > 0.3$). Regarding the
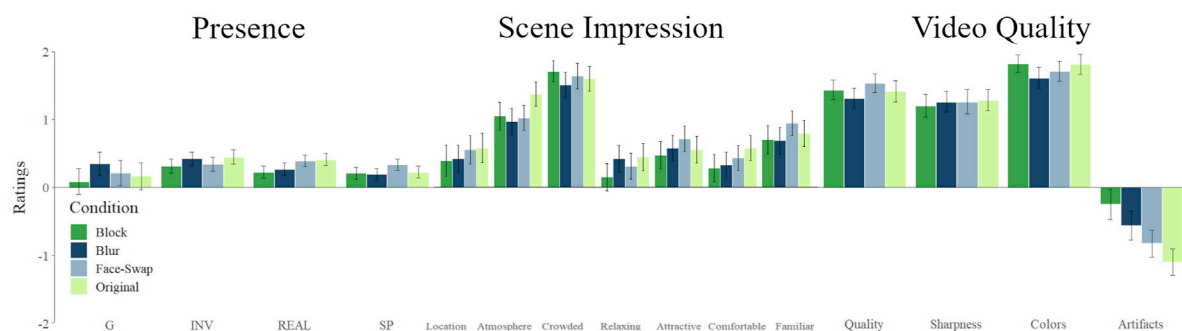
FIGURE 2
Barplot visualizing the answers for the questionnaires displayed after each trial for the presence questionnaire (G - General Presence, INV - Involvement, REAL - Experienced Realism, SP - Spatial Presence), scene impression, and video quality.
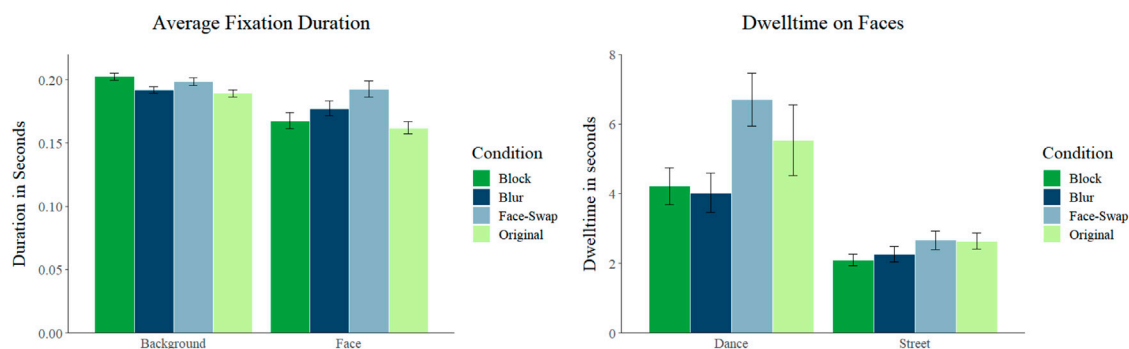

FIGURE 3
Barplots visualizing the eye tracking data. Left: Average duration of fixations for fixations on the background and faces, Right: Overall dwell time on faces for dance and street videos.

video quality questionnaire, we find that participants perceive the occurrence of artifacts differently between the conditions (ANOVA $F(3, 316) = 3.1$, $p < 0.03$, $\eta^2 = 0.03$) with *post hoc* Tukey HSD tests highlighting differences between videos anonymized using blocking and non-anonymized videos ($[-1.6, -0.9]$ $p < 0.02$). If we separate the stimuli to consider only on videos that focus on humans as in the dance parade (see Figure 1), we can observe a tendency of lower realism rating and worse reports of scene impressions for blocked videos but do not detect significant differences in the analysis.

3.2 Eye tracking data

To evaluate the eye tracking data, we estimate fixations and the amount and speed of saccades from the raw data to compare the conditions (Engbert and Kliegl, 2003). We find that, there are no differences between the amount (ANOVA $F(3, 630) = 0.14$, $p > 0.05$) or speed (ANOVA $F(3, 26432) = 1.34$, $p > 0.05$) of saccades. However, we do detect differences in the length of fixations between the conditions (ANOVA $F(3, 26432) = 6.97$, $p < 0.005$, $\eta^2 = 0.008$). Post hoc Tukey HSD tests indicate differences between the conditions Block and Non-Anonymized ($[-0.02, -0.003]$ $p < 0.005$) as well as Face-swap and Non-Anonymized ($[-0.023, -0.005]$ $p < 0.0005$).

As facial anonymization changes only the facial areas of videos, we furthermore perform investigations in changes of viewing behavior for fixations on facial areas (Wöhler et al., 2024). Here, we define the facial area based on the face detection results which were used to apply the facial anonymization techniques. This way, fixations on these areas do not only correctly match faces in the videos but also correctly indicate the area blurred or blocked by anonymization. By classifying each fixation as either looking at a face or not, we can understand whether facial anonymization affects the attention to faces. Looking at the average fixation duration again from this perspective, as visualized in Figure 3 (Left), we see that blocked videos have longer fixations on the background, while in face-swapped videos the fixations on faces are longer than in the non-anonymized videos. Next, we compare the overall dwell time on faces between the conditions. A statistical assessment shows significant differences between the overall dwell times on faces between the conditions (ANOVA $F(3, 26432) = 20.98$, $p < 0.005$, $\eta^2 = 0.02$). In Figure 3 (Right) the dwell time on faces are shown for two videos of a dance parade as well as the general street scenes. It is visible that the dance parade leads to higher dwell times on faces while in street scenes participants focus more on the background.

3.3 Memorization of scene content

After viewing all videos, participants were presented with still frames and asked to decide whether they saw the content during the experiment. To analyze the data, we look into whether the ability of participants to correctly answer the question was affected by the anonymization technique used in the video corresponding to the frame. We find that participants were best at correctly assessing frames for videos shown in Face-Swap 90.00%, followed by Block 83.33%, Blur 79.31%, and finally non-Anonymized videos 64.84%. Despite the drop in correct assessments, we do not find significant differences between the conditions (ANOVA F (3, 236) = 1.5, p > 0.05).

3.4 Debriefing

In the debriefing, we asked participants which facial anonymization they noticed and to describe their impression of the techniques. Additionally, we asked whether they noticed face-swapping and also had them describe their opinion of the technology.

We find that all participants noticed blocking, while 76% noticed blurring and only 30% noticed face-swapping. Blocking was found to be distracting by six participants and three reported reduced realism of the videos. Among the participants that noticed blurring, two reported it to be distracting and to reduce the realism of the videos. Among the participants that noticed face-swapping, one reported to not have a strong opinion of the technique as they rather focused on other video elements. In contrast, four participants described a feeling of discomfort due to factors like artifacts, people having the same face, and faces and bodies that are not matching each other. Finally, one participant described face-swapping as generally unnatural.

4 Discussion

4.1 H1: presence and scene impression

In our experiment, participants were not informed about the investigation of facial anonymization. As the stimuli mostly consist of street videos, participants tend to focus more on the surrounding scenery than faces. Due to this, blurring and face-swapping were only seldom reported to be distracting in the debriefing. We did not find significant differences between the sense of presence or scene impression between the conditions in our statistical assessment. Therefore, we do not find evidence for hypothesis H1. This indicates that facial anonymization may not have a negative impact for videos focusing on street scenes in free viewing tasks. As the results may change for scenes focusing on faces, further studies with stimuli focusing on human performances or dialogues would be interesting.

4.2 H2: video quality

We assessed negative effects of facial anonymization on the video quality. We find that participants report more artifacts for videos anonymized by blocking the facial area than for non-anonymized videos. This indicates that the black areas are perceived as artifacts. We think this is mainly due to flickering that occurs when face detection results change between frames, which results in appearing and disappearing of the black areas between frames. Such artifacts were reported in previous work in which participants described the flickering as distracting (Wöhler et al., 2024). In contrast, the quality of videos that use blurring or face-swapping as anonymization does not differ significantly from real videos. Therefore, we find only partial support for our hypothesis H2.

4.3 H3: eye tracking

Based on the eye tracking data, we observe differences in gazing behavior and visual attention between the conditions. We find that the duration of fixations is significantly shorter for non-anonymized videos than for face-swapped and blocked videos. Moreover, we find significant differences between the overall dwell time on faces between the conditions. The results highlight that participants spend more time looking at faces for face-swapped and non-anonymized videos. Therefore, we find support for hypothesis H3.

4.4 H4: memory of scene content

After the experiment, we showed images to participants and asked them to identify whether each image belongs to a video they previously saw or not. We do not find significant differences between the answers of each condition and find no support for hypothesis H4. This indicates that facial anonymization in street videos does not necessarily affect the ability of participants to memorize the scene. However, videos focusing on educational or touristic purposes that aim to convey specific knowledge to the viewer and also use audio might be differently affected by facial anonymization. As previous work found various effects on learning outcomes based on the presence and visibility of instructors (Beege et al., 2025), it would be interesting to see how the anonymization of instructors in 360° videos affect learners.

4.5 Comparison to previous work

In contrast to a previous study (Wöhler et al., 2024), we did not find differences in consideration of presence and scene perception between the conditions. As previous work informed participants about the focus on facial anonymization, participants might have deliberately inspected the faces, which could have more strongly impacted their ratings. As our experiment does not instruct participants about facial anonymization and has a free viewing task, our results may be more similar to natural viewing scenarios than the previous study. Therefore, we conclude that the choice of facial anonymization might be less relevant for the feeling of presence in street videos if it is assumed the attention of participants lies more on the environment than faces. However, videos focusing on humans like videos of a dance parade could profit from face-swapping or no anonymization.

4.6 Quality of face-swapping

We find that only a small number of participants recognized face-swapping, further adding evidence to the unobtrusiveness of the technology (Wöhler et al., 2021; Khamis et al., 2022). However, in cases where participants noticed the application of face-swapping, they report artifacts focusing on a feeling of discomfort (Wöhler et al., 2024). In our study, we use the framework SimSwap (Chen et al., 2020), which we chose due to its ability to efficiently generate face-swaps that can handle small occlusions (e.g., glasses). As previous work suggests that the ability of participants to distinguish between face-swaps and real faces worsens for more modern approaches (Bozkir et al., 2024), more recent face-swapping algorithms might further reduce the perception of artifacts leading to even less obtrusive facial anonymization. For the practical application, it would be valuable to consider more modern approaches as well as the trade-off between quality and processing efficiency.

4.7 Implication of the eye tracking results

Our evaluation shows that the attention of participants is affected by facial anonymization techniques. We find that face-swapped and non-anonymized videos have longer dwell times on faces. This seems intuitive as blocking and blurring completely remove the facial details and leave nothing to see for the viewer. We also find that non-anonymized videos have a different average fixation duration than blocked and face-swapped videos which could indicate that the reported flickering artifacts introduced by facial anonymization attract the gaze of viewers and disrupt fixations. This is in line with previous works that use flickering in peripheral areas to guide the gaze of viewers in 360° videos viewed in HMD (Schmitz et al., 2020). As face-swapped videos tend to have longer fixation durations on faces, it is possible that participants investigate faces after noticing artifacts or a feeling of discomfort that were reported in the debriefing.

4.8 Limitations

The main limitation of our study is the limited number of participants. As we recruit 20 participants, power analysis indicates that we can only detect significant differences (0.05) with medium effect sizes (Cohen's d = 0.2) with a power of 86% for the questionnaires. Therefore, our negative results regarding the scene presence, scene impression, and memorization might be due to the limited sample size. Additionally, all our participants are Japanese and the anonymized videos are taken in Japan which means that we cannot assess cultural differences or familiarity effects which are especially interesting for facial modifications, e.g., due to the same-race bias (Chiroro and Valentine, 1995).

4.9 Opportunities for future work

As gazing behavior and therefore the visual attention of participants is impacted by facial anonymization, more studies focusing on videos with different purposes are necessary. As the stimuli used in this paper focused on showcasing city scenes, participants spent most of the viewing time inspecting the background. Therefore, the effect of facial anonymization could be even more pronounced in videos focusing on humans, e.g., educational videos with a guide for virtual field trips.

Furthermore, it would be valuable to assess how methods that promise even stronger anonymization, e.g., by anonymizing the gait and voice of recorded people are perceived (Hanisch et al.,2025). Such anonymization tools might be especially valuable for viewing private data like medical records in VR.

Finally, recent work investigates privacy protection for augmented reality devices using blocking of facial areas (Corbett et al., 2023). It would be interesting to apply face-swapping to see if it can increase the perceived realism and allow the preservation of facial expressions. However, distracting artifacts in the generated faces could be even more severe in augmented scenarios as they could detract users from the real environment increasing the chance for accidents.

5 Conclusion

Our results suggest that the choice of facial anonymization technique in 360° videos is relevant for the visual attention of viewers and the perception of the video quality. Especially the gathered eye tracking data show that participants focus more on the surroundings and less on people for videos anonymized with traditional techniques. While face-swapping shows more similar overall dwell times on faces, the average fixation duration is longer than in non-anonymized videos. This could mean that participants fixate longer on each face to inspect it for artifacts. We do not detect differences between the facial anonymization techniques in presence and memorization. This indicates that anonymization might not generally have a negative impact on videos that do not focus on faces and instead introduce the surrounding area like tourism videos.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the Graduate School of Information Science and Technology, The University of Tokyo. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LW: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. SI: Conceptualization, Investigation, Supervision,

Writing – original draft, Writing – review and editing. KA: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Beege, M., Schroeder, N. L., Heidig, S., Rey, G. D., and Schneider, S. (2025). The instructor presence effect and its moderators in instructional video: a series of meta-analyses. *Educ. Res. Rev.* 41, 100564. doi:10.1016/j.edurev.2023.100564

Bonnail, E., Frommel, J., Lecolinet, E., Huron, S., and Gugenheimer, J. (2024). "Was it real or virtual? confirming the occurrence and explaining causes of memory source confusion between reality and virtual reality," in *Proceedings of the ACM human factors in computing systems (CHI)* (New York, NY, USA: Association for Computing Machinery), 1–17. doi:10.1145/3613904.3641992

Bozkir, E., Riedmiller, C., Skodras, A. N., Kasneci, G., and Kasneci, E. (2024). Can you tell real from fake face images? perception of computer-generated faces by humans. *ACM Trans. Appl. Percept.* 22, 1–23. doi:10.1145/3696667

Chao, Y.-P., Kang, C.-J., Chuang, H.-H., Hsieh, M.-J., Chang, Y.-C., Kuo, T. B., et al. (2023). Comparison of the effect of 360 versus two-dimensional virtual reality video on history taking and physical examination skills learning among undergraduate medical students: a randomized controlled trial. *Virtual Real.* 27, 637–650. doi:10.1007/s10055-022-00664-0

Chen, R., Chen, X., Ni, B., and Ge, Y. (2020). "Simswap: an efficient framework for high fidelity face swapping," in *Proceedings of the ACM international conference on multimedia* (New York, NY, USA: Association for Computing Machinery), 1–9. doi:10.1145/3394171.3413630

Chiroro, P., and Valentine, T. (1995). An investigation of the contact hypothesis of the own-race bias in face recognition. *Q. J. Exp. Psychol.* 48, 879–894. doi:10.1080/14640749508401421

Corbett, M., David-John, B., Shang, J., Hu, Y. C., and Ji, B. (2023). "Bystandar: protecting bystander visual data in augmented reality systems," in *Pro. International conference on Mobile systems, applications and services* (New York, NY, USA: Association for Computing Machinery), 370–382. doi:10.1145/3581791.3596830

Engbert, R., and Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vis. Res.* 43, 1035–1045. doi:10.1016/s0042-6989(03)00084-1

Faklaris, C., Cafaro, F., Blevins, A., O'Haver, M. A., and Singhal, N. (2020). A snapshot of bystander attitudes about mobile live-streaming video in public settings. *Informatics* 7, 10–15. doi:10.3390/informatics7020010

Hanisch, S., Arias-Cabarcos, P., Parra-Arnau, J., and Strufe, T. (2025). Anonymization techniques for behavioral biometric data: a survey. *ACM Comput. Surv.* 57, 3729418. doi:10.1145/3729418

Hasan, E. T., Hasan, R., Shaffer, P., Crandall, D., Kapadia, A., et al. (2017). "Cartooning for enhanced privacy in lifelogging and streaming videos," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (New York, NY, USA: IEEE), 29–38.

Khamis, M., Farzand, H., Mumm, M., and Marky, K. (2022). "Deepfakes for privacy: investigating the effectiveness of state-of-the-art privacy-enhancing face obfuscation methods," in *Proceedings of the international conference on advanced visual interfaces* (New York, NY, USA: Association for Computing Machinery), 1–5. doi:10.1145/3531073.3531125

Khamis, M., Panskus, R., Farzand, H., Mumm, M., Macdonald, S., and Marky, K. (2024). "Perspectives on deepfakes for privacy: comparing perceptions of photo owners and obfuscated individuals towards deepfake versus traditional privacy-enhancing obfuscation," in *Proceedings of the international conference on Mobile and ubiquitous multimedia* (New York, NY, USA: Association for Computing Machinery), 300–312. doi:10.1145/3701571.3701602

Kumar, K., Poretski, L., Li, J., and Tang, A. (2022). Tourgether360: collaborative exploration of 360 videos using pseudo-spatial navigation. *Proc. ACM Human-Computer Interact.* 6, 1–27. doi:10.1145/3555604

Li, Y., Vishwamitra, N., Knijnenburg, B. P., Hu, H., and Caine, K. (2017). Effectiveness and users' experience of obfuscation as a privacy-enhancing technology for sharing photos. *Proc. ACM Human-Computer Interact.* 1, 1–24. doi:10.1145/3134702

Liu, R., Xu, X., Yang, H., Li, Z., and Huang, G. (2022). Impacts of cues on learning and attention in immersive 360-degree video: an eye-tracking study. *Front. Psychol.* 12, 792069. doi:10.3389/fpsyg.2021.792069

Mizuho, T., Narumi, T., and Kuzuoka, H. (2024). "Investigating the effects of changing the appearance of screen-based avatars on audience memory," in *ACM symposium on applied perception 2024* (New York, NY, USA: Association for Computing Machinery), 1–9.

Regenbrecht, H., and Schubert, T. (2002). Real and illusory interactions enhance presence in virtual environments. *Presence Teleoperators Virtual Environ.* 11, 425–434. doi:10.1162/105474602760204318

Schmitz, A., MacQuarrie, A., Julier, S., Binetti, N., and Steed, A. (2020). "Directing versus attracting attention: exploring the effectiveness of central and peripheral cues in panoramic videos," in *2020 IEEE conference on virtual reality and 3D user interfaces (VR)* (New York, NY, USA: IEEE), 63–72.

Schubert, T. W. (2003). The sense of presence in virtual environments: a three-component scale measuring spatial presence, involvement, and realness. *Z. für Medien.* 15, 69–71. doi:10.1026//1617-6383.15.2.69

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence Teleoperators Virtual Environ.* 10, 266–281. doi:10.1162/105474601300343603

Takenawa, M., Sugimoto, N., Wöhler, L., Ikehata, S., and Aizawa, K. (2023). "360rvw: fusing real 360° videos and interactive virtual worlds," in *Proceedings of the ACM international conference on multimedia* (New York, NY, USA: Association for Computing Machinery), 9379–9381. doi:10.1145/3581783.3612670

Thomas Ruberto, A. D. A., Chris, M., and Semken, S. (2023). Comparison of in-person and virtual grand canyon undergraduate field trip learning outcomes. *J. Geoscience Educ.* 71, 445–461. doi:10.1080/10899995.2023.2186067

Wilson, E., Shic, F., Skytta, J., and Jain, E. (2022). Practical digital disguises: leveraging face swaps to protect patient privacy. *arXiv Prepr. arXiv:2204.03559 abs/2204.03559.* doi:10.48550/arXiv.2204.03559

Wöhler, L., Zembaty, M., Castillo, S., and Magnor, M. (2021). "Towards understanding perceptual differences between genuine and face-swapped videos," in *Proceedings of the ACM human factors in computing systems (CHI)* (New York, NY, USA: Association for Computing Machinery), 1–13. doi:10.1145/3411764.3445627

Wöhler, L., Ikehata, S., and Aizawa, K. (2024). Investigating the perception of facial anonymization techniques in 360° videos. *ACM Trans. Appl. Percept. (TAP)* 21, 1–17. doi:10.1145/3695254