Face-Swapped Privacy: Understanding Preferences for Video Obfuscation of Oneself in Mobile Contexts

Marvin Strauss
University of Duisburg-Essen
Essen, Germany
marvin.strauss@uni-due.de

Jonas Keppel
University of Duisburg-Essen, Essen, Germany
Essen, Germany
jonas.keppel@uni-due.de

Jan Günther
University of Duisburg-Essen
Essen, Germany
jan.guenther.1412@stud.uni-due.de

Verena Winterhalter
LMU Munich
Munich, Germany
verena.winterhalter@ifi.lmu.de

Donald Degraen
University of Canterbury
Christchurch, New Zealand
donald.degraen@gmail.com

Viktorija Paneva
LMU Munich
Munich, Germany
viktorija.paneva@ifi.lmu.de

Uwe Gruenefeld
GenerIO
Essen, Germany
uwe@generio.ai

Jonas Auda
GenerIO
Essen, Germany
jonas@generio.ai

David Goedicke
University of Duisburg-Essen
Essen, Germany
david.goedicke@uni-due.de

Florian Alt
LMU Munich
Munich, Germany
florian.alt@ifi.lmu.de

Stefan Schneegass
University of Duisburg-Essen, Essen, Germany
Essen, Germany
stefan.schneegass@uni-due.de



(a) A Masked Face in a Doctor's Waiting Room

(b) A Blurred Face while Jogging in the Park

(c) A Face-Swapped Face in a Living Room

Figure 1: We investigated video face obfuscation through an online pre-study that selected contrasting recording contexts and a laboratory study that compared masking, blurring, and face-swapping applied to participants' own and another person's face. Participants perceived masking as the strongest method, although the methods did not differ in their overall ranking. They prioritized stronger protection for themselves and an enjoyable viewing experience when evaluating another person's obfuscation. Context further shaped comfort with recording and the perceived need for obfuscation.

Abstract

Prior work on privacy-preserving face replacement suggests that face-swapping can protect identity while preserving the visual experience of static images. We examine whether these conclusions extend to a narrower setting: face obfuscation in short, contextually framed videos that include participants' own faces. Through an online study (N = 121) and a laboratory study (N = 18), we

This work is licensed under a Creative Commons Attribution 4.0 International License.

MUM '26, Glasgow, United Kingdom

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2791-7/2026/12
https://doi.org/10.1145/3861188.3861193

compare masking, blurring, and face-swapping across recording contexts and between self- and other-obfuscation. Participants perceived masking as stronger than face-swapping for their own videos and reported greater discomfort with face-swapping, while evaluating face-swapping more favorably when it was applied to another person. Context influenced comfort with being recorded and the perceived need for obfuscation. These findings qualify, rather than causally contradict, prior image-based evidence: conclusions about generative replacement do not uniformly transfer to the self-obfuscation scenarios studied here. We derive implications for context-sensitive and preference-aware video obfuscation.

CCS Concepts

• **Security and privacy → Privacy protections**; **Usability in security and privacy**.

Keywords

Privacy, Anonymization, Obfuscation, Face-Swap, Usability, User Study

ACM Reference Format:

Marvin Strauss, Jonas Keppel, Jan Günther, Verena Winterhalter, Donald Degraen, Viktorija Paneva, Uwe Gruenefeld, Jonas Auda, David Goedicke, Florian Alt, and Stefan Schneegass. 2026. Face-Swapped Privacy: Understanding Preferences for Video Obfuscation of Oneself in Mobile Contexts. In *25th International Conference on Mobile and Ubiquitous Multimedia (MUM '26), December 01–04, 2026, Glasgow, United Kingdom.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3861188.3861193

1 Introduction

Cameras are ubiquitous in everyday life, and bystanders may be recorded unintentionally by smartphones, webcams, or camera-equipped Extended Reality (XR) devices [4, 31–34, 36]. Because bystanders often cannot determine whether and what a device is recording [7], footage containing them may be shared online without their awareness or consent [38]. Such recordings can reveal sensitive information and thereby create privacy risks [25].

A common approach to protecting recorded individuals is to make their faces unrecognizable [3]. Established methods include masking (Figure 1a) and blurring (Figure 1b) [46]. Although effective, these methods are visually conspicuous and can reduce the aesthetic quality of the footage [16]. Generative approaches instead replace privacy-sensitive content with synthetic alternatives [45]. Face-swapping, for example, replaces the original face with a new one (Figure 1c) while preserving the overall visual coherence of the content [42, 45].

Prior work suggests that face-swapping can provide effective and aesthetically pleasing obfuscation in static images [16]. Individuals also perceive their own faces as protected when face-swapping is applied to personal photos [17]. Whether these findings extend to videos remains unclear. Facial motion can facilitate recognition, particularly for familiar faces [19, 20, 30], and face-swapping typically preserves the original facial movements [42]. Consequently, the preservation of facial motion may affect how face-swapping is perceived in videos, revealing a potential trade-off between privacy protection and visual appeal.

Privacy perceptions are also strongly context-dependent [29]. People may tolerate weaker but less disruptive obfuscation in public settings, whereas recordings in more private or sensitive environments may require stronger protection. Accordingly, the suitability of an obfuscation method may depend not only on its visual properties, but also on where a recording occurs and whose face is being protected.

We investigate these factors through an online study and a laboratory study. The online study ($N = 121$) examines how people assess the sensitivity of different recording contexts and their need for obfuscation. The laboratory study ($N = 18$) evaluates masking, blurring, and face-swapping applied to participants' own faces and to another person's face in videos representing contexts of different sensitivities (Figure 1).

Our findings show that participants perceived masking as stronger than face-swapping and generally prioritized stronger protection for their own faces over visual appeal. In contrast, they favored more aesthetically pleasing obfuscation when another person was shown. We further find that recording context influences participants' comfort and perceived need for obfuscation. These results indicate that face-swapping is not a universally suitable solution for video obfuscation and motivate context-sensitive controls that account for the preferences of the person being protected.

Our work builds directly on human-centered evaluations of generative image protection for static images [16, 45]. We therefore do not claim that evaluating subjective perceptions of generative obfuscation is itself novel. Instead, we test the applicability of prior image-based conclusions in a distinct setting in which the protected content is specifically facial identity, participants encounter their own familiar faces, the stimuli are short contextually framed videos, and judgments are made from both the protected person's and a viewer's perspective.

Contribution Statement. We contribute (1) an empirical comparison of masking, blurring, and face-swapping for obfuscating one's own and another person's face in contextually situated videos; (2) evidence from an online study ($N = 121$) and a laboratory study ($N = 18$) on how context and the person being obfuscated shape preferences for privacy protection and visual appeal; and (3) design implications for context-sensitive and preference-aware video obfuscation.

2 Related Work

This work draws from several strands of prior research, most importantly, obfuscation methods and the contextual influence on privacy and image recording.

2.1 Face Obfuscation Methods and Perceptions of One's Own Face

Masking and blurring are widely used to protect privacy-sensitive facial information in images and videos [5, 15, 21, 23, 24, 48]. However, both methods can reduce visual quality, and blurring may be vulnerable to deblurring or machine-based re-identification [22, 28, 39].

Generative approaches such as face-swapping offer an alternative to conventional obfuscation by replacing the original face with

a synthetic identity while attempting to preserve the visual coherence of the content. Although deepfakes are commonly associated with deception and other security risks [18, 41, 43], recent work has explored their use for privacy protection. For example, DeepPrivacy removes the detected facial region and generates a new face conditioned on privacy-safe information such as sparse pose and the surrounding image context [13]. Face-swapping can therefore conceal facial identity without introducing the conspicuous visual disruptions associated with masking or blurring.

Human-centered studies have examined how such generative face obfuscation is perceived in static images. Khamis et al. [16] compared synthetic face replacement with masking, blurring, pixelation, and avatars and found that it could protect facial identity while preserving the visual aesthetics of photographs. Xu et al. [45] investigated a related but broader approach called generative content replacement. Rather than focusing specifically on faces, their study covered 23 categories of privacy-sensitive content in third-party dataset images and compared GCR with blurring, cartooning, color filling, and removal. They found that GCR exhibited low edit detectability, hindered the identification of specific content, and achieved high visual harmony and overall satisfaction. Xu et al. thus establish broader perceptual advantages of generative replacement, whereas our work focuses specifically on face-swapping and facial identity.

The perspective of the person whose face is being obfuscated has also been examined in static photographs. Khamis et al. [17] asked participants to apply obfuscation methods to themselves and others in their personal photos. Both photo owners and depicted individuals perceived face-swapping as effective and visually well integrated. Together, these studies show that generative face replacement can balance identity protection and visual experience in static images. However, they leave open how these conclusions apply when face-swapping is used in short videos of participants' own familiar faces, when facial and head movements remain visible, and when evaluations consider recording context and differences between self- and other-obfuscation.

2.2 Challenges in Obfuscating Video Footage

Applying face-swapping to videos introduces challenges that are less pronounced in static images. Dynamic facial features support recognition [12, 20, 40] and remain largely preserved because face-swapping transfers the original facial movements to the synthetic face [42]. Such motion cues can facilitate recognition even under adverse viewing conditions and are particularly important for familiar faces [11, 19, 30].

Initial work on video obfuscation reflects this concern. Wöhler et al. [44] compared blurring, black boxes, and face-swapping in 360° videos. Although face-swapping was perceived as the most realistic and least disruptive method, participants questioned its effectiveness for protecting identity. However, it remains unclear how people perceive face-swapping when it is applied to their own faces in videos and how recording context shapes this assessment.

2.3 Contextual Influence on Privacy and Image Recording

Privacy perceptions are context-dependent and shaped by situational norms governing who shares what information with whom and under which conditions [2, 29]. Violating these context-specific expectations can constitute a privacy breach.

Recent work on everyday AR sensor use further identified recurring categories of sensitive situations, including the presence of confidential information, expectations of solitude, home environments, and public locations [1]. Such classifications motivate covering a broad range of situations, but do not determine which concrete video-recording scenarios provide useful contrasts for evaluating face obfuscation. Our pre-study therefore serves the narrower purpose of selecting contextual configurations for the subsequent laboratory study.

Prior work has therefore explored context-aware privacy mechanisms for image recording and sharing. These approaches adapt obfuscation to crowd density [9], allow individuals to specify context-dependent protection preferences [35], or adjust sharing decisions and information granularity based on image content, user behavior, and requester context [49]. However, it remains unclear how recording context influences people's perceptions of the suitability of different face obfuscation methods.

3 Research Approach

Building on image-based research on face-swapping and generative content replacement [16, 17, 45], we investigate how these findings apply to face obfuscation in videos. Rather than benchmarking the general perceptual properties of generative replacement, we focus on how recording context and the person being protected shape perceptions of masking, blurring, and face-swapping. Compared with Khamis et al.'s [16] initial viewer-oriented evaluation of face obfuscation in static images, we examine participants' own familiar faces as well as another person's face. We further extend their subsequent study of self- and other-obfuscation in personal photographs [17] by using videos across systematically selected recording contexts. This face-specific focus also distinguishes our research questions from Xu et al.'s [45] broader evaluation of GCR across diverse privacy-sensitive content in dataset-based static images.

We address the following research questions:

RQ1 How does recording context shape perceptions of face obfuscation?
RQ2 Which face obfuscation methods do people prefer for their own face in videos?
RQ3 How do people perceive obfuscation methods when someone else is obfuscated in videos?

We conducted two studies. An online study examined the perceived sensitivity of different recording contexts and informed the selection of three contexts for the laboratory study. In the laboratory study, participants evaluated videos of themselves and another person across these contexts, obfuscated using masking, blurring, and face-swapping. Together, the studies address the influence of context, obfuscation method, and the person being obfuscated.

Ethics. Both studies received approval from our institutional ethics committee. All participants were informed about the procedure and data protection practices and provided informed consent.

4 Pre-Study: Evaluating the Sensitivity of Being Recorded in Different Contexts

To answer **RQ1** and select contrasting recording contexts for the subsequent laboratory study, we conducted an online pre-study. The purpose was not to derive a comprehensive taxonomy of sensitive recording situations, but to empirically calibrate a bounded set of candidate scenarios for studying face obfuscation in videos.

4.1 Study Design

Our online pre-study used a within-subject design with Context as the independent variable. We constructed ten candidate scenarios using a maximum-variation approach grounded in contextual integrity and prior work on context-dependent privacy perceptions [1, 10, 29]. The scenarios covered home and institutional settings, routine public activities, and situations that could disclose health-related, political, or potentially stigmatizing information. In selecting them, we varied the physical setting, bystander activity, likely recording actors and audiences, potentially disclosed information, and expected information flows. Thus, context was treated as more than location: each scenario represented a particular configuration of social interactions, activities, audiences, and possible disclosures. The scenarios therefore covered a diverse set of privacy-relevant recording situations, as shown in Figure 2. The videos were sourced from Pexels[1].

4.2 Measurements

For each scenario, participants answered four Likert items adapted from Hadan et al. [10], assessing recording likelihood, comfort, information sensitivity, and the need for obfuscation. Two open-ended questions elicited reasons for participants' (dis)comfort and possible mitigation strategies. We additionally measured general online privacy concerns using the IUIPC scale [26] to characterize the sample. IUIPC scores and demographic variables were collected to characterize the sample and were not included as predictors in our inferential analyses. The complete questionnaire is provided in Table 1 in the appendix.

4.3 Procedure and Recruiting

We conducted the study in LimeSurvey[2] and recruited participants through Prolific[3]. Participants received £3.80, and the median completion time was 28 minutes. After providing study information, participants viewed the ten contexts in randomized order, each presented through a short description and video, and completed the corresponding questions. Two attention checks were included.

4.4 Participants

Recruiting via Prolific, we obtained a total of $N = 121$ participants who completed our online survey. Among them, 61 identified as women, 59 as men, and one as non-binary. Thereby, our participants originated from Europe (67), Africa (35), as well as from North and South America, Asia, and Australia (19). They were aged between 18 and 82 ($M = 34.32$, $SD = 13.19$). The participants' IUIPC scores had a mean of 5.57 ($SD = 1.18$) for *Control*, 5.25 ($SD = 1.17$) for *Collection*, and 6.00 ($SD = 1.14$) for *Awareness*.

4.5 Results

We analyzed the four context-dependent Likert items using Friedman tests, followed by Bonferroni-corrected Wilcoxon signed-rank tests. Descriptive distributions are provided in Figure 9 in the appendix, together with the complete post-hoc comparisons in Figure 10.

Context significantly affected perceived recording likelihood ($\chi^2(9) = 126.05, p < .001$), comfort with being recorded ($\chi^2(9) = 136.09, p < .001$), need for obfuscation ($\chi^2(9) = 115.81, p < .001$), and sensitivity of disclosed information ($\chi^2(9) = 186.05, p < .001$). Recordings at a *Red Light* and a *Demonstration* were perceived as particularly likely, although several public contexts, including *Shopping* and *Jogging*, were also rated as more likely than indoor settings such as the *Waiting Room*, *Living Room*, and *Office*. The *Waiting Room* and *Red Light* consistently elicited the strongest privacy concerns: participants reported comparatively low comfort, a high need for obfuscation, and high sensitivity of the information disclosed. The *Demonstration* was likewise perceived as revealing relatively sensitive information, whereas the *Living Room* was generally associated with lower recording likelihood, less need for obfuscation, and lower perceived information sensitivity.

4.6 Implications for Further Investigation

To select contexts for the laboratory study, we prioritized comfort with being recorded, need for obfuscation, and sensitivity of disclosed information, while perceived recording likelihood served as a realism check. We sought lower-, intermediate-, and higher-sensitivity configurations rather than selecting only the scenarios that elicited the strongest privacy concerns. This allowed the laboratory study to examine face obfuscation across contrasting levels of perceived sensitivity and different activity structures. Based on these indicators, we selected *Living Room*, *Jogging*, and *Waiting Room* as lower-, intermediate-, and higher-sensitivity contexts.

Living Room represented the lower-sensitivity condition because participants generally reported greater comfort, less need for obfuscation, and lower information sensitivity. Qualitative responses further associated this setting with familiar and trusted recorders. *Jogging* served as the intermediate condition because it occupied neither sensitivity extreme and introduced an open, public, and physically active setting that contrasted with the two seated indoor contexts. *Waiting Room* represented the higher-sensitivity condition due to comparatively low comfort, a strong need for obfuscation, and concerns about revealing health-related information.

These contexts provide analytically useful contrasts for the laboratory study rather than representative proxies for all recording situations with similar levels of sensitivity.

4.7 Limitations

The pre-study was designed to select contrasting scenarios for the laboratory study, not to derive a representative taxonomy or estimate the prevalence of privacy concerns across everyday recording

[1] https://www.pexels.com, last retrieved November 5, 2026.
[2] https://www.limesurvey.org, last retrieved November 5, 2026.
[3] https://www.prolific.com, last retrieved November 5, 2026

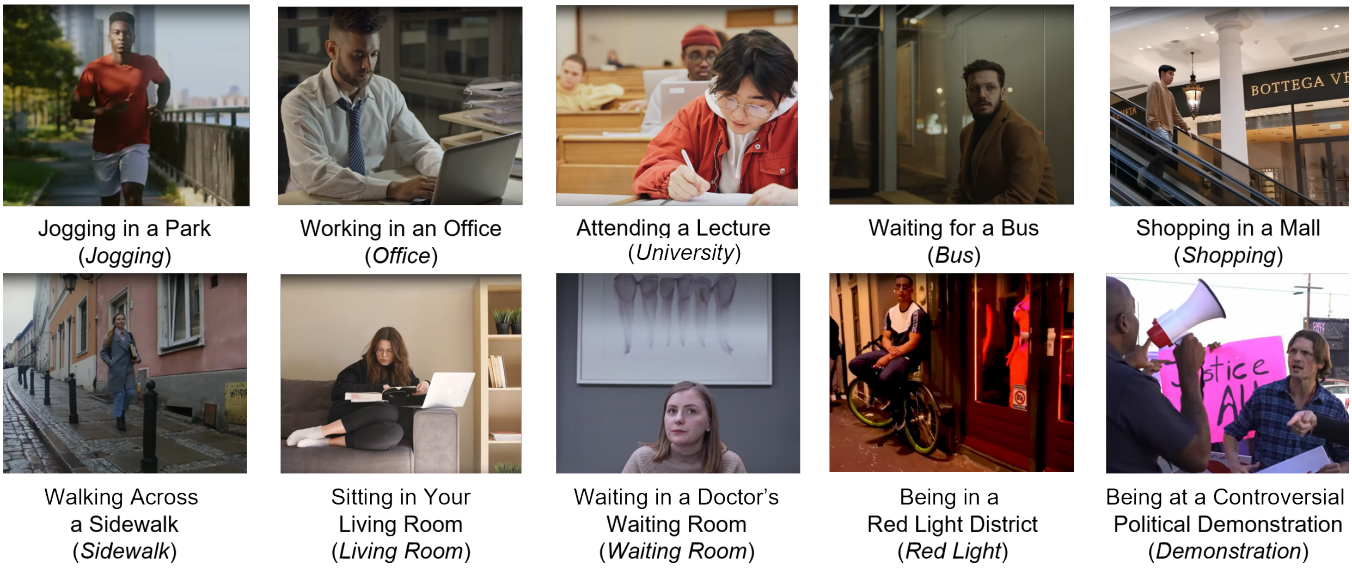


Figure 2: Ten different contexts in which a bystander is recorded are visualized through frames of videos that participants viewed before rating each context based on its sensitivity and the perceived need for obfuscation.

situations. The ten stock-video scenarios were purposively selected, and each combined several contextual properties, including setting, activity, likely recording actors and audiences, and potentially disclosed information. Ratings therefore apply to the presented configurations, and omitted situations or alternative descriptions of the same settings might have produced a different ordering. Accordingly, the three selected contexts should be understood as experimental anchors rather than as representative categories of low-, intermediate-, and high-sensitivity recording situations.

5 Laboratory Study: Comparing Perceptions of Different Obfuscation Methods

Building on the insights from our online pre-study, we evaluate how bystanders perceive different obfuscation methods in videos when they are accidentally recorded by a camera in various contexts to answer our research questions.

5.1 Study Design

Our lab study uses a within-subject design. We distinguish three independent variables: Context, Method, and Person. The Context variable includes the three selected settings with varying levels of sensitivity, as detailed in Subsection 4.6: *Waiting Room*, *Jogging*, and *Living Room*. Method comprises three obfuscation techniques, which we refer to as *Masking*, *Blurring*, and *Face-Swap*. Person has two levels: *Own* and *Other's* obfuscation.

5.2 Apparatus

To evaluate participants' perceptions of being recorded and obfuscated across contexts, we captured a facial image and recorded three approximately 15-second videos of each participant in front of a well-lit green screen using a full-HD webcam. Participants sat in an armchair to represent a *Living Room*, sat on a chair to represent a doctor's *Waiting Room*, and jogged in place for the *Jogging* context.

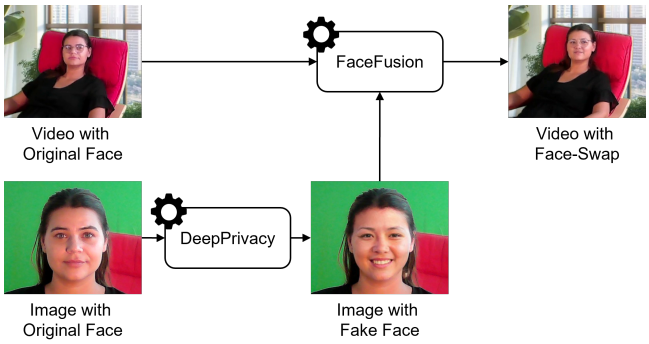


Figure 3: Face-Swapping Process: DeepPrivacy first generates a synthetic face from an image containing the bystander's face. FaceFusion then uses this synthetic face, along with the original face video, to produce a face-swapped video. The participant gave her consent to the publication of the pictures taken.

We recorded the videos using Open Broadcaster Software (OBS)[4], which replaced the green screen with context-specific backgrounds. We used static images for the *Living Room* and *Waiting Room* and a reversed video of a park passing by to simulate running outdoors[5].

We applied all three Methods to each video. For *Face-Swap*, we followed Khamis et al. [16] and used DeepPrivacy [13] to generate one synthetic source face for each participant. DeepPrivacy replaces the detected facial region using a conditional generative model that

[4]https://obsproject.com, last retrieved November 5, 2026.
[5]https://www.pexels.com/video/an-off-road-cutting-trough-the-forest-5594981/, last retrieved November 5, 2026.

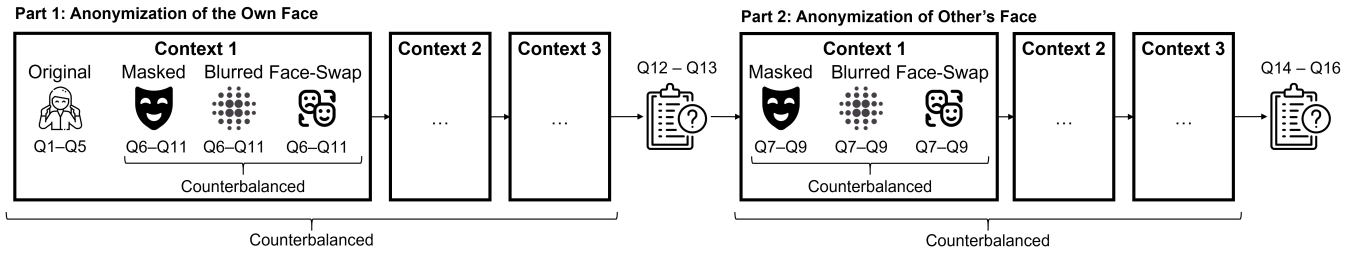


Figure 4: Procedure of the Lab Study: In the first part, participants viewed videos of themselves in a counterbalanced order across different contexts. Each context began with a non-obfuscated video, followed by videos with different types of face obfuscation, with corresponding questions. Afterward, they answered general questions about face obfuscation. In the second part, participants were shown obfuscated videos of another person and asked corresponding questions, also in a counterbalanced order. Finally, they responded to general questions about the obfuscation of other persons.

draws on privacy-safe information such as sparse facial pose and the surrounding image context rather than the original facial pixels. We then used the synthetic face and the original video as input to FaceFusion[6] to produce the face-swapped video. The process is illustrated in Figure 3.

For *Masking* and *Blurring*, we used YOLOv8[7] to detect faces in each frame and Dlib[8] to identify facial contours. We processed the detected regions using OpenCV[9] and Pillow[10], following parameters used in prior work [16, 24]. For blurring, we applied Pillow's `GaussianBlur()` with a radius proportional to the frame dimensions; frames of 770×552 pixels used a radius of 4. For masking, we set all pixels within the predicted facial contour to black. Example outputs are shown in Figure 1a and Figure 1b.

5.3 Measurements

Participants provided demographic information and completed the IUIPC questionnaire to characterize the sample. These measures were not used for subgroup comparisons or as predictors. To assess perceptions of being recorded across contexts, we showed participants the videos with their original faces and asked them to imagine being accidentally recorded by another person. Using 7-point Likert scales, they rated recording likelihood, comfort, need for obfuscation, and sensitivity of the disclosed information, and explained their (dis)comfort in an open-ended response. We used 7-point scales throughout the laboratory study for consistency.

For each obfuscated video, participants rated six aspects: ease of self-recognition, the range of people who might still recognize them, perceived obfuscation strength, suitability for the context, satisfaction with the edited video, and discomfort with the facial modification. The recognizability item was inspired by Dunbar's social-circle model [8] and ranged from recognition by nobody to recognition even by strangers. Because self-recognition and recognition by familiar others may not directly correspond to the general perception of obfuscation strength, we measured perceived strength separately. For example, some participants might consider an obfuscation effective if it prevents recognition by strangers, whereas others may only consider it effective if they cannot recognize themselves. Following Xu et al. [45], we measured overall satisfaction with the edited video. We treat satisfaction as a broad evaluative judgment rather than as a direct measure of aesthetic quality. Ratings may reflect visual appearance, recognizability, discomfort with the replacement identity, or other reactions to the edit. We did not separately assess naturalness, visual harmony, or edit detectability. Our remaining measures were not direct replications of Xu et al.'s constructs: whereas they assessed viewers' ability to infer original content, we examined perceived obfuscation strength and recognizability from the protected person's perspective.

Participants additionally ranked the three methods for obfuscating their own faces and explained their choices. To compare self-obfuscation with the obfuscation of others, they evaluated nine pre-recorded videos featuring a male third person across the three contexts and methods. For these videos, they rated satisfaction, obfuscation strength, and contextual suitability. Finally, we assessed whether participants prioritized strong protection or an enjoyable viewing experience for their own and another person's obfuscation using two Likert items and an open-ended question. The complete set of measures is provided in Table 2, Table 3, and Table 4.

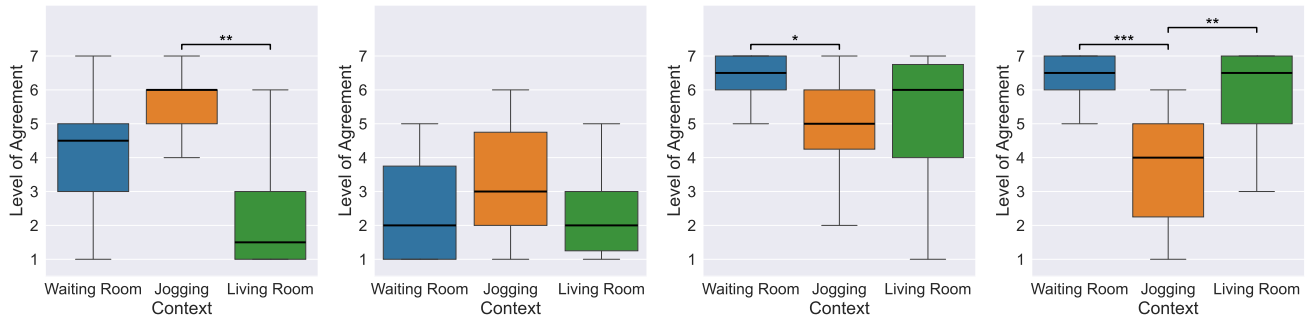
5.4 Procedure

After receiving a study description and providing informed consent, participants recorded a facial image and the three context videos. While the obfuscated videos were processed, participants received further information about the study and completed the demographic questions and IUIPC questionnaire.

In the first part, participants evaluated videos of themselves context by context. For each context, they first viewed the original video, followed by the three obfuscated versions, and completed the corresponding measures. The order of contexts and obfuscation methods was counterbalanced using a Latin square. Participants then answered general questions about the three methods.

In the second part, participants evaluated videos of a third person across all contexts and obfuscation methods. The obfuscated and original versions were presented side by side, with the obfuscated face on the left. Contexts and methods were again counterbalanced.

[6] https://docs.facefusion.io/, last retrieved November 5, 2026.
[7] https://yolov8.com/, last retrieved November 5, 2026.
[8] http://dlib.net/, last retrieved November 5, 2026.
[9] https://opencv.org/, last retrieved November 5, 2026.
[10] https://pillow.readthedocs.io, last retrieved November 5, 2026.



(a) Recording Likelihood (b) Comfort with Recording (c) Need for Obfuscation (d) Sensitivity of Information

Figure 5: Boxplots visualizing the ratings for the different Likert items regarding the sensitivity of being recorded in a context on a 7-Point Likert Scale.

Finally, participants reflected on differences between obfuscating themselves and another person. The full procedure is illustrated in Figure 4.

5.5 Participants

We recruited a total of $N = 18$ participants for our lab study through university email lists and on-campus outreach. Of these participants, 9 identified as women and 9 as men, with ages ranging from 21 to 32 years ($M = 26.17$, $SD = 3.17$). The participants' IUIPC scores had a mean of 5.93 ($SD = .75$) for *Control*, 6.06 ($SD = .86$) for *Collection*, and 6.35 ($SD = .75$) for *Awareness*.

5.6 Limitations

Our laboratory stimuli were short, controlled green-screen videos that differed not only in contextual meaning but also in visual and behavioral properties. The *Jogging* condition involved physical activity, an upright posture, and a moving background, whereas the *Waiting Room* and *Living Room* conditions involved seated participants and static backgrounds. Although the videos contained facial and bodily motion, they did not include evolving interactions, activities, or environmental changes and therefore represent constrained contextual configurations rather than the full range of naturalistic video. Consequently, differences between conditions cannot be attributed exclusively to contextual sensitivity, and differences from prior image-based findings cannot be attributed specifically to facial motion or the video medium. Posture, participant motion, background dynamics, and other differences between studies may also have influenced the ratings.

Participants were also composited into researcher-selected contexts rather than recorded in situations they personally encountered or selected. Consequently, they had to imagine the recording scenario, and the selected contexts cannot represent the diversity of individual privacy expectations. Moreover, context extends beyond physical setting to include the recorder, intended audience, purpose, information flow, and subsequent use of the footage. Our results should therefore be interpreted as comparisons between experimentally controlled contextual configurations rather than general preferences for broad categories such as public or private spaces.

The laboratory sample was small ($N = 18$) and covered a limited age range. It was therefore not suitable for estimating demographic effects or stable participant-level differences. The qualitative responses illustrate variation within this sample but should not be interpreted as indicating the prevalence of particular preferences in a broader population.

Finally, our findings are specific to the evaluated pipeline and one DeepPrivacy-generated face per participant. Because FaceFusion preserves pose and facial motion while leaving cues such as hair and head shape visible, perceived similarity, recognition, discomfort, and satisfaction may reflect both the generated identity and remaining cues. Different synthetic faces or pipelines could therefore yield different results.

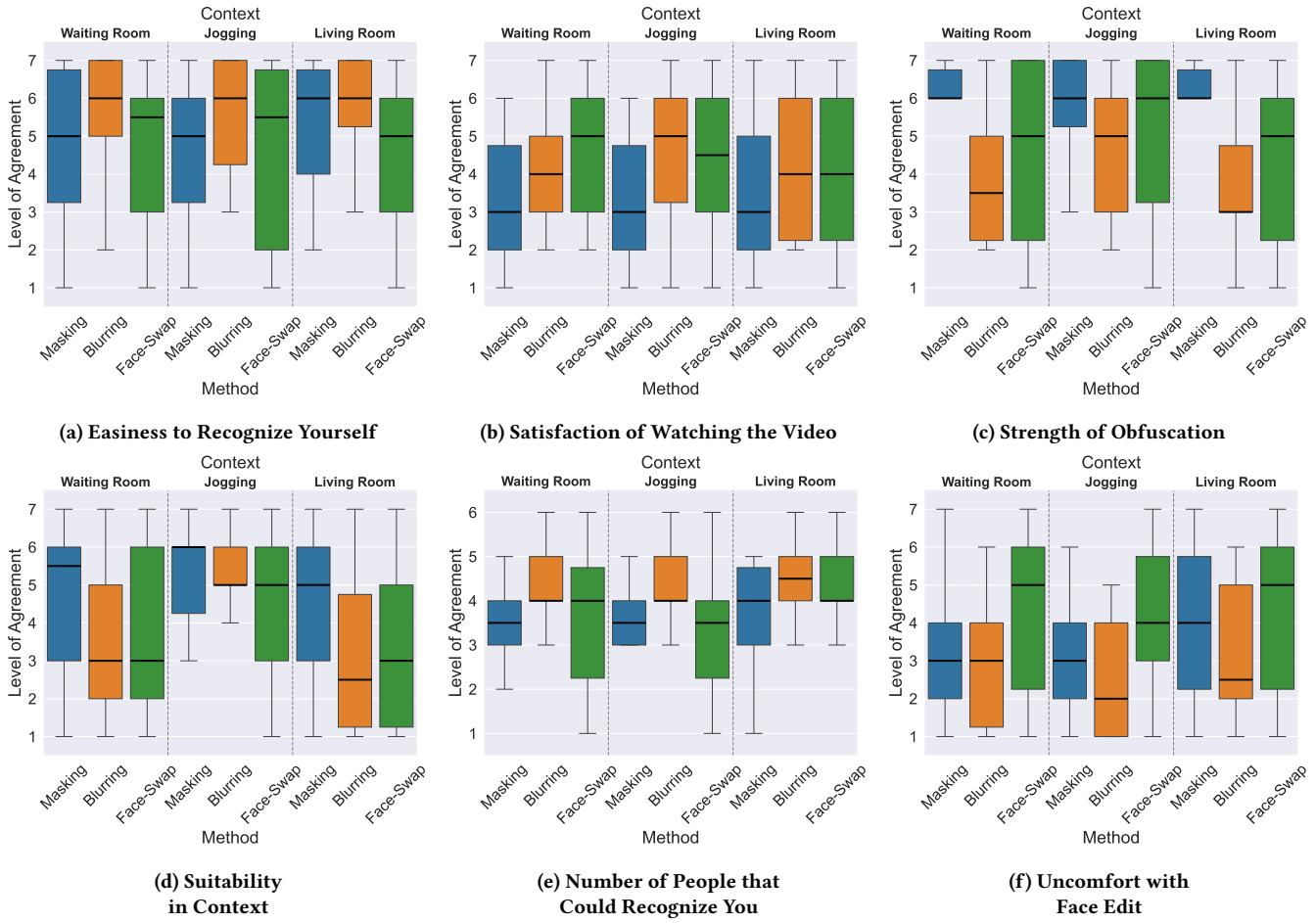
5.7 Results

In the following sections, we present the results of our lab study. The first three sections focus on the quantitative findings related to participants' perceptions of being recorded in different contexts, their views on their own obfuscated videos, and their reactions to the obfuscation of others. Finally, we perform a thematic analysis to explore the qualitative feedback provided by the participants.

5.7.1 Being Recorded in Different Contexts. We analyzed the four context-related Likert items using Friedman tests, followed by Bonferroni-corrected Wilcoxon signed-rank tests. Descriptive results are shown in Figure 5.

Context significantly affected perceived recording likelihood ($\chi^2(2) = 11.2, p = .003$). Participants considered accidental recording more likely while *Jogging* than in their *Living Room* ($p = .003$); the remaining comparisons were not significant. We found no significant effect on comfort with being recorded ($\chi^2(2) = 4.41, p = .11$).

Context also affected the perceived need for obfuscation ($\chi^2(2) = 8.98, p = .001$), which was higher in the *Waiting Room* than while *Jogging* ($p = .019$). Finally, perceived information sensitivity differed across contexts ($\chi^2(2) = 21.03, p < .001$): *Jogging* was rated as less sensitive than both the *Waiting Room* ($p < .001$) and *Living Room* ($p = .002$), whereas the two indoor contexts did not differ significantly.



(a) Easiness to Recognize Yourself

(b) Satisfaction of Watching the Video

(c) Strength of Obfuscation

(d) Suitability in Context

(e) Number of People that Could Recognize You

(f) Uncomfort with Face Edit

Figure 6: Boxplots illustrating the ratings for the different Likert items related to the perception of various obfuscation methods applied to one's own face across different contexts.

5.7.2 Self-Obfuscation in Different Contexts. We analyzed participants' ratings of their own obfuscated videos using Aligned Rank Transform ANOVAs with Method and Context as independent variables. Significant effects were followed by Bonferroni-adjusted estimated marginal mean comparisons. Descriptive results are shown in Figure 6.

Method affected the ease of self-recognition ($F(2, 153) = 3.86, p = .023$). Participants found themselves easier to recognize when *blurred* than when *face-swapped* ($t(157) = 3.37, p = .003$). The remaining method comparisons were not significant. Neither Context nor the interaction affected self-recognition.

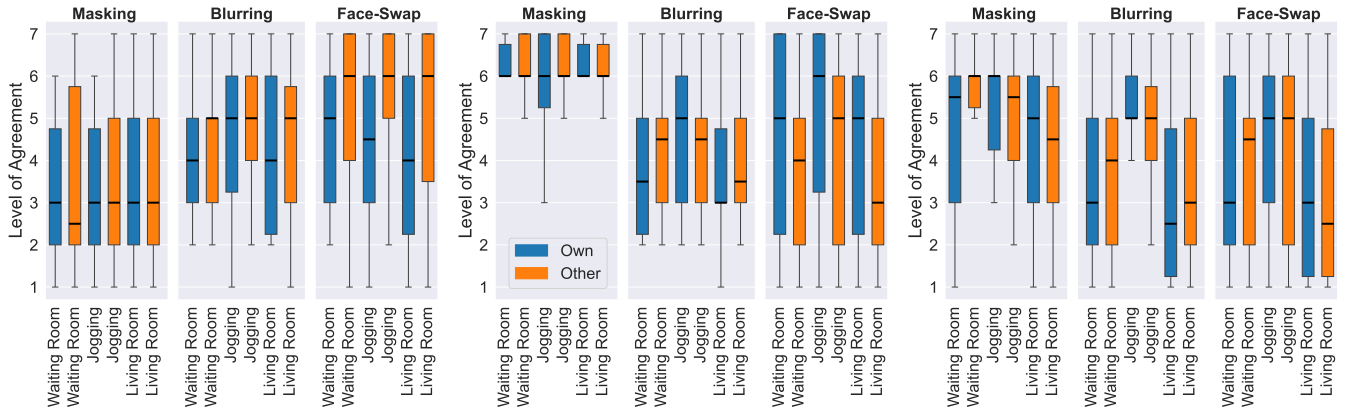
Method also affected satisfaction with the edited videos ($F(2, 153) = 3.95, p = .021$). Participants were more satisfied with *blurring* than with *masking* ($t(157) = 2.56, p = .033$), while the other comparisons were not significant. We again found no effects of Context or interaction.

Perceived obfuscation strength differed substantially between methods ($F(2, 153) = 14.67, p < .001$). *Masking* was perceived as stronger than both *blurring* ($t(157) = 5.87, p < .001$) and *face-swapping* ($t(157) = 3.90, p < .001$), whereas *blurring* and *face-swapping* did not differ. Context and the interaction were not significant.

For contextual suitability, we found effects of both Method ($F(2, 153) = 4.10, p = .019$) and Context ($F(2, 153) = 7.80, p < .001$). *Masking* was considered more suitable than *blurring* ($t(157) = 2.44, p = .048$) and *face-swapping* ($t(157) = 2.63, p = .028$). Across methods, obfuscation was rated as more suitable while *Jogging* than in the *Waiting Room* ($t(157) = 2.63, p = .028$) or *Living Room* ($t(157) = 3.46, p = .002$). We found no interaction effect.

Method further affected the perceived number of people who could recognize the participant ($F(2, 153) = 4.96, p = .008$). Participants believed that more people would recognize them when *blurred* than when *masked* ($t(157) = 3.22, p = .005$) or *face-swapped* ($t(157) = 2.48, p = .043$). Context and the interaction were not significant.

Finally, discomfort with the edited face differed by Method ($F(2, 153) = 6.52, p = .002$). Participants felt more uncomfortable with *face-swapping* than with *blurring* ($t(157) = 3.58, p = .002$). The



(a) Satisfaction of Watching the Video

(b) Strength of Obfuscation

(c) Suitability in Context

Figure 7: Boxplots illustrating the ratings for the various Likert items related to satisfaction, strength, and suitability of different obfuscation methods for oneself (blue) and others (orange) across various contexts.

remaining comparisons were not significant. No effects of Context or interaction emerged.

Participants' overall ranking of the three methods did not differ significantly ($\chi^2(2) = 2.33, p = .311$). Descriptive results are presented in Figure 8a.

Taken together, the measures reveal differentiated profiles rather than a single superior method. Masking was perceived as the strongest and most contextually suitable method. Participants expected fewer people to recognize them when masked or face-swapped than when blurred, although they found themselves easier to recognize when blurred than when face-swapped. For satisfaction, only the difference between blurring and masking was significant; face-swapping did not differ from either alternative. Face-swapping also produced greater discomfort than blurring. Context affected contextual suitability but not the other measures, and no interactions between Context and Method emerged. Finally, participants' overall ranking of the methods did not differ significantly.

5.7.3 Self-Obfuscation and Obfuscation of Others in Different Contexts. To compare perceptions of self-obfuscation and the obfuscation of others, we analyzed satisfaction, perceived strength, and contextual suitability separately for each Method using ART ANOVAs with Context and Person as independent variables. Significant effects were followed by Bonferroni-adjusted estimated marginal mean comparisons. Descriptive results are shown in Figure 7.

For satisfaction, neither Context nor Person affected ratings for *Masking* or *Blurring*. For *Face-Swap*, however, we found a significant effect of Person ($F(1, 102) = 7.87, p = .006$): participants reported greater satisfaction with face-swapping when it was applied to another person than to themselves ($t(104) = 2.38, p = .019$). No Context or interaction effects emerged.

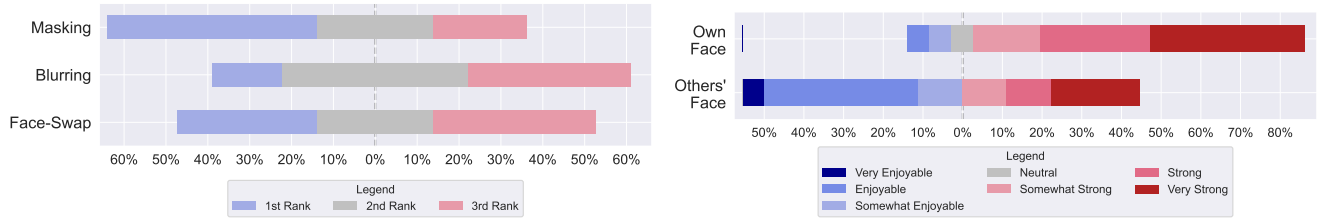
Perceived obfuscation strength did not differ significantly by Context, Person, or their interaction for any of the three methods. For contextual suitability, we found no significant effects for *Masking* or *Face-Swap*. In contrast, suitability ratings for *Blurring* differed by Context ($F(2, 102) = 8.38, p < .001$): blurring was considered more suitable while *Jogging* than in the *Waiting Room* ($t(104) = 3.27, p = .004$) or *Living Room* ($t(104) = 3.95, p < .001$). The two indoor contexts did not differ, and no Person or interaction effects emerged.

Finally, a Friedman test showed that preferences for strong versus enjoyable obfuscation differed depending on the person being protected ($\chi^2(1) = 9, p = .003$). Participants preferred stronger obfuscation for their own faces, but more visually enjoyable obfuscation for others. The corresponding results are shown in Figure 8b.

5.7.4 Thematic Analysis. We conducted a thematic analysis [6] of the open-ended responses. Two researchers independently open-coded a subset of responses ($N = 7$), discussed their codes to develop an initial coding tree, and applied it to the full dataset. Remaining disagreements were resolved through discussion, resulting in three themes.

Contextual Influence. Participants emphasized that the suitability of obfuscation depends on the recording context. As P9 stated: *"There may be situations where [...] I don't need to be obfuscated. On the other hand, if I'm being filmed in a context that makes me uncomfortable [...] I would like to be completely obfuscated as much as possible."* Comfort was shaped by the personal information revealed and the inferences others could draw. P3, for example, expressed concern about the living room: *"I don't want to find any recordings of it anywhere, as the living room is usually decorated in a very personal way."* Similarly, P12 noted that *"The setting of this doctor's room alone allows [...] conclusions to be drawn about my illnesses."*

Other participants regarded some contexts as less sensitive or had become accustomed to being recorded. P5 commented about jogging: *"It's out in the woods, and anyone could see it anyway. [...] If there are cameras, then it's just normal you get filmed."* Responses were not uniform. While many considered the living room sensitive, P9 stated: *"I don't feel uncomfortable because the person recording me has already been invited into my living room."*



(a) Ranking of the Face Obfuscation Methods

(b) Strong vs. Enjoyable Obfuscation

Figure 8: Participants' ranking of the three obfuscation methods and their preference for a strong or enjoyable obfuscation for themselves and others.

Face Obfuscation Methods. Face-swapping was often perceived as visually appealing but insufficiently protective because participants still recognized themselves. P12 stated: *"The change to my face was too close to me. My hair stayed almost the same; the shape of my head was exactly the same."* Others valued its subtlety, as P18 noted: *"You don't assume that this face is covering another person."* However, this subtlety also raised concerns about deception. P17 argued: *"You should be able to see that there is actually another face behind it."*

Participants also described face-swapping as aesthetically pleasing but sometimes unsettling. P8 commented: *"It is a human face and therefore looks the most beautiful to me,"* whereas P15 stated: *"It's just a bit weird to see yourself like that."* P16 additionally questioned the origin of the replacement face: *"The face will come from somewhere and then someone else's face will be taken and that's not cool."*

Masking was viewed as strong but visually disruptive. P3 explained: *"I find it more suitable to block out the face completely because then it is the strongest method available,"* while P5 remarked: *"This black box just looks awful."* Some participants also noted that masking shifted attention to other identifying features. P10 stated: *"I somehow looked at the whole body, which somehow broke this obfuscation for me, it helps me to focus on the person."*

Blurring was often seen as a compromise, although participants questioned its effectiveness. P12 noted: *"All of the facial features were immediately recognizable,"* and P8 added: *"I'm not sure if the blurring can be resolved."* In contrast, P17 described it as a middle ground: *"I think you can still see a bit of my face, but you don't recognize me."*

Participants further argued that face-only obfuscation may be insufficient in sensitive settings. As P6 stated: *"So then it's even better to somehow fade out the people completely and not just the face."*

Difference Between Own and Others' Obfuscation. Participants prioritized strong privacy protection for their own footage. P14 emphasized: *"The main thing is that I am obfuscated. It's not about whether others find it enjoyable."* P3 similarly stated: *"I personally would like to protect my data, especially when it comes to biometric markers, like the face."* At the same time, some participants cared about their appearance, as P5 noted: *"I don't want to look so weird."*

For others' obfuscation, participants more often prioritized aesthetics. P12 remarked: *"When I watch something that I find exciting, I don't want everything to be full of black boxes, because it's distracting. But this face-swap would look cool."* Others questioned this asymmetry. P8 stated: *"If it's important to me personally that people don't recognize me, it should also apply to all other people."* P11 summarized the underlying tension: *"It's a double-edged sword, at the same time I don't want to be recognized, but it should still look nice."*

6 Discussion

Our evaluations reveal contextual differences in perceptions of video obfuscation, variation across participants, and tensions between obfuscating one's own face versus others. We discuss these findings in relation to our research questions and provide design recommendations for developers and video creators. The following recommendations translate participants' perceptions into support for user decision-making. They should not be interpreted as technical security guarantees or as validated rules for automatically selecting an obfuscation method. Our study assessed perceived strength and recognizability, not actual resistance to human or automated identification.

6.1 Contextual Influence on Video Obfuscation

In addressing **RQ1**, both studies found that participants' ratings of comfort, information sensitivity, and the need for obfuscation differed across the presented contextual configurations. These differences appear closely tied to the type of information revealed in each context. For instance, recordings in a doctor's waiting room may disclose potential health conditions, whereas footage from a living room could reveal personal preferences.

Moreover, the individual recording the footage and the intended audience influence privacy perceptions. While sharing videos with friends and family is typically not viewed as intrusive, privacy concerns arise when companies exploit such recordings for purposes like targeted advertising. These findings align with Nissenbaum's framework of contextual integrity [29], which emphasizes that both the nature of the information and the recipient's identity are critical factors in shaping the flow of personal information and influencing privacy perceptions.

In the laboratory study, ratings differed across the three contextual configurations. Across methods, obfuscation was considered

more suitable in the *Jogging* condition than in the two seated indoor conditions. Because these stimuli varied jointly in contextual meaning, posture, physical activity, and background motion, we interpret this as a difference between the presented configurations rather than as an isolated effect of contextual sensitivity. Moreover, the absence of an interaction between Context and Method means that our results do not establish that a particular obfuscation method is preferred for a particular context. Instead, the findings motivate controls that allow users to adjust protection to the circumstances and their own preferences.

Participants' qualitative responses were not uniform. For example, while some described jogging in public as harmless, others expressed concerns about their physical appearance during the activity. The online and laboratory studies also produced different assessments of the living-room scenario. Such variation is consistent with prior research showing that privacy perceptions can differ across individuals and situations [37, 47].

However, our studies did not test whether demographics, stable privacy preferences, or sample composition explained this variation. The two studies also differed in their stimuli, procedures, response scales, and whether participants viewed themselves in the footage. We therefore cannot attribute the different assessments of the living-room scenario to demographic or individual differences and treat these only as possible interpretations rather than findings supported by our analyses.

Together with prior evidence on the individual nature of privacy perceptions, the variation in participants' responses motivates configurable privacy controls rather than uniform or demographically inferred defaults. Systems should allow the people being protected to adjust obfuscation to their own preferences and circumstances.

Recommendation 1: Elicit Context and Audience Rather Than Infer Them

Before a video is shared, interfaces should ask the creator and, where possible, the depicted person to indicate the intended audience, privacy concerns, and desired degree of concealment. These inputs should remain editable and should be used to organize obfuscation alternatives rather than to infer privacy needs automatically from the scene or demographic information.

6.2 Preferences for One's Own Obfuscation in Videos

Regarding **RQ2**, participants perceived masking as the strongest method for protecting their own faces, whereas face-swapping was perceived as less protective and sometimes left recognizable cues in participants' appearance, head shape, hair, or movements. Importantly, we found no evidence that face-swapping provided a satisfaction advantage over masking or blurring in the self-obfuscation condition, and it produced greater discomfort than blurring. Satisfaction should therefore not be interpreted as an isolated measure of aesthetic quality. Instead, it represents a broader evaluation that may incorporate visual appearance, self-recognition, and discomfort with the replacement identity.

These measures capture distinct judgments and were analyzed separately. We therefore cannot determine whether participants who were more satisfied with a face-swap also found themselves more recognizable or whether discomfort contributed to their satisfaction ratings. The qualitative responses nevertheless illustrate the underlying tension: some participants described face-swapping as visually appealing and less disruptive than masking, while others found their replacement identity unsettling. Thus, face-swapping may preserve visual subtlety without necessarily being perceived as sufficiently protective for one's own familiar identity. Examining the within-participant relationships among these judgments requires a larger sample and an explicitly exploratory multilevel analysis. We position this result within the broader face-swapping literature in Subsection 6.4.

Furthermore, our findings raise ethical questions about the visibility of obfuscation, especially face-swapping. Should obfuscation be intentionally detectable to avoid misleading viewers? This issue has already been explored in the context of deepfakes [18] and is reinforced by our qualitative data. The literature suggests that increased privacy awareness can influence user behavior [27], and in the case of face obfuscation, making viewers aware of obfuscation techniques might encourage more responsible sharing and use of obfuscated footage.

The particular synthetic identity generated for each participant formed part of the experienced face-swapping condition. Although DeepPrivacy generates facial content without access to the original facial pixels, the subsequent FaceFusion step and the unchanged parts of the video preserve other identity-related cues, including hair, head shape, body shape, clothing, posture, and facial and head movements. Participants' perceived similarity to their face-swapped representation can therefore not be attributed solely to the generated facial features. Moreover, similarity and comfort need not vary in the same direction: a more dissimilar synthetic identity might reduce self-recognition but could also intensify discomfort with being represented by an unfamiliar face. Our study therefore does not identify an optimal degree of similarity; it shows how participants responded to the particular synthetic identities and generation pipeline used here.

Participants' observations that face-only obfuscation leaves such cues visible also motivate an escalation step in which users can extend the edited region after previewing the complete video. Possible options include full-body masking or replacement, cropping the person from the frame, or concealing identifying elements in the background. For instance, DeepPrivacy 2 [14] enables realistic full-body replacement in images and may provide broader obfuscation while preserving visual coherence. However, our study did not evaluate the effectiveness of these alternatives. Their implementation and validation remain future work.

Recommendation 2: Preview Alternatives and Support Escalation

Masking was perceived as the strongest method, but the methods did not differ in their overall ranking. Interfaces should therefore preview masking, blurring, and face-swapping in the intended video rather than declare a universal winner. For face-swapping, users could compare several synthetic identities. If identifying cues remain in the hair, head shape, body, clothing, movement, or background, they should be able to extend obfuscation beyond the face. Masking may be presented as a conservative option when perceived protection is prioritized.

6.3 Difference Between Own and Others' Obfuscation in Videos

Regarding **RQ3**, participants placed greater emphasis on protection strength when considering their own faces and on an enjoyable viewing experience when considering another person's face. Masking was rated as the strongest method, but the overall ranking of the three methods did not differ significantly. We therefore interpret this result as a difference in the criteria participants prioritized rather than as evidence that they universally preferred masking for themselves.

For face-swapping specifically, participants reported greater satisfaction when it was applied to another person than when it was applied to themselves. The qualitative responses similarly suggested that participants were more willing to accept visually subtle obfuscation for others, while some questioned whether this asymmetry was ethically appropriate.

This finding extends research on the perspectives of photo owners and depicted individuals in static photographs [17] by directly contrasting the same participants' evaluations of self- and other-obfuscation in videos. It suggests that viewer evaluations may not fully represent the preferences of the person whose identity is being protected.

Recommendation 3: Give the Depicted Person Control Over Approval

A creator's preference for an enjoyable edit should not override the protection preferences of the depicted person. Where feasible, interfaces should provide the depicted person with a preview of the edited clip and request approval for the selected method, particularly before face-swapping is applied. If creator and depicted person disagree, sharing should pause until they agree on an edit or a mutually accepted, more conservative alternative is selected.

6.4 Positioning the Findings Within Prior Face-Obfuscation Research

Consistent Observations. Several observations align with earlier research on generative privacy protection. Participants described face-swapping as visually appealing and valued its ability to blend into the footage without the conspicuous disruption caused by masking. This is consistent with Khamis et al.'s [16] findings that synthetic face replacement can preserve the aesthetics of static photographs. It also conceptually aligns with the low edit detectability and high visual harmony reported for the broader generative content replacement approach studied by Xu et al. [45], as well as with Wöhler et al.'s [44] observation that face-swapping appeared realistic and less disruptive in 360° videos. Because our study did not directly measure edit detectability, narrative coherence, or visual harmony, these similarities should be understood as consistent observations rather than direct replications of those measures.

Qualifications of Prior Findings. Our results qualify the conditions under which the benefits of generative face replacement apply. Khamis et al. [17] found face-swapping to be perceived as effective and visually well integrated when participants applied it to themselves and others in personal static photographs. In our short video stimuli, participants instead perceived masking as stronger for protecting their own identities and frequently reported self-recognition and discomfort with their face-swapped appearance. Similarly, we did not observe a general satisfaction advantage for face-swapping when it was applied to participants' own videos. These findings do not contradict the broader potential of generative replacement demonstrated by Xu et al. [45]. Rather, they show that its reported perceptual advantages do not uniformly transfer to face-swapping in the self-obfuscation scenarios studied here.

Additional Insights. Our study further contributes evidence about factors that were not jointly examined in this prior work. First, evaluations differed between self- and other-obfuscation: participants prioritized stronger protection for themselves but evaluated face-swapping more favorably when another person was shown. Second, recording context shaped participants' comfort with being recorded, the perceived sensitivity of the disclosed information, and their need for obfuscation. Together, these findings show that the suitability of face-swapping cannot be determined solely from the perspective of a viewer or from the visual properties of the edited footage. It must also account for the preferences of the person being protected and the circumstances in which the recording occurs.

Possible Explanations for Differences Between Images and Videos. Videos introduce temporal identity cues that are absent from static images. In our stimuli, face-swapping preserved participants' facial expressions and head movements, which may have supported self-recognition despite the replacement of static facial features [19, 20, 30, 42]. Temporal presentation may also make inconsistencies or unnatural transitions in the generated face more noticeable than they would be in a single image. These video-specific properties provide plausible explanations for why participants perceived face-swapping as less protective or more unsettling than reported in prior image-based studies.

However, our study was not designed to determine which of these properties caused the observed differences. Facial motion and temporal presentation varied alongside personal familiarity, generation quality, task framing, and participant population. Our findings therefore show that conclusions from static-image studies do not uniformly transfer to the self-obfuscation scenarios examined here, but they do not establish whether motion, temporal artifacts, or another difference between the studies was primarily responsible.

Recommendation 4: Account for Differences Between Image and Video Obfuscation

Videos introduce temporal cues that are absent from static images. In face-swapped footage, preserved facial and head movements as well as potential inconsistencies across frames may affect recognizability and discomfort. Developers should not assume that image-based evaluations transfer unchanged to video. Video obfuscation should additionally be evaluated for motion-based recognizability and temporal consistency.

6.5 Toward a User-Controlled Obfuscation Workflow

Taken together, our findings motivate a user-controlled workflow rather than automatic selection of a supposedly optimal method. First, the creator specifies the intended audience and sharing purpose, while the depicted person is invited to state relevant concerns and the desired degree of concealment. Second, the interface presents masking, blurring, and face-swapping as previews within the actual video rather than as isolated frames. The previews should communicate that the methods differ in perceived strength, satisfaction, recognizability, and discomfort, without presenting these perceptions as guarantees of actual protection.

Third, the depicted person selects or approves an edit. If their preference conflicts with that of the creator, sharing should pause until a mutually acceptable edit is found. Fourth, the person reviews the complete edited video for remaining cues in the face, body, movement, clothing, or environment and can escalate from face-only editing to broader obfuscation. A future semi-automated system could use the elicited concerns to order or suggest alternatives, but our findings do not validate an algorithm for automatically selecting a method.

This workflow may be particularly relevant for smartglasses and other wearable cameras, where prolonged or unobtrusive recording can capture bystanders who have little control over the resulting footage [1, 7, 36]. In such applications, the wearer may favor a visually unobtrusive edit, while the depicted person may prioritize stronger perceived protection—a conflict reflected in the asymmetry observed in our study. These possibilities are application directions rather than features validated by our studies. Field research is needed to determine how preference elicitation and approval can operate when bystanders cannot readily be contacted.

7 Conclusion and Future Work

This paper investigated how recording context and obfuscation method shape perceptions of masking, blurring, and face-swapping when applied to one's own and another person's face in videos. Based on our findings, we provide design recommendations for developers and video creators toward more context-sensitive and preference-aware obfuscation.

Our results show that context influences comfort with being recorded, the perceived need for obfuscation, and the suitability of obfuscation methods. For self-obfuscation, masking was consistently perceived as the strongest method, whereas face-swapping was considered less protective because participants still recognized aspects of their identity and sometimes felt uncomfortable with their altered appearance. Face-swapping was evaluated more favorably when applied to another person, revealing an asymmetry between preferences for protecting oneself and judgments made as a viewer.

These findings both support and qualify prior research. Consistent with Khamis et al. [16], qualitative responses highlighted the visual appeal and subtlety of synthetic face replacement. However, participants' responses to their own face-swapped videos differ from findings in personal static photographs, where face-swapping was perceived as effective and visually well integrated [17]. They further show that the broader perceptual advantages of generative content replacement reported by Xu et al. [45] do not uniformly transfer to the face-specific self-obfuscation scenarios studied here. Preserved facial and head movements offer a plausible explanation, as motion can support the recognition of familiar faces. However, because our design does not isolate motion or medium from familiarity, generation quality, and task framing, we do not attribute these differences to video alone.

Our findings emphasize that privacy-preserving systems should involve the people whose identities are being protected rather than relying solely on the preferences of viewers or content owners. At the same time, face-swapping and other forms of realistic content replacement retain potential for balancing identity protection and visual experience. Future work should investigate different synthetic identities and generation pipelines, systematically examine the role of facial motion, and evaluate more naturalistic recordings. It should also explore whether full-body replacement can provide stronger protection while preserving visual coherence. Overall, obfuscation should be personalized and sensitive to the protected individual, recording context, intended audience, and desired level of protection.

Acknowledgments

This work is funded by the German Research Foundation (DFG) under project no. 521584224 and is part of Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.

References

[1] Melvin Abraham, Mohamed Khamis, and Mark McGill. 2024. Don't Record My Private pARts: Understanding The Role of Sensitive Contexts and Privacy Perceptions in Influencing Attitudes Towards Everyday Augmented Reality Sensor Usage. In *2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 749–758. https://doi.org/10.1109/ISMAR62088.2024.00090

[2] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. 2015. Privacy and human behavior in the age of information. *Science* 347, 6221 (2015), 509–514. https://doi.org/10.1126/science.aaa1465

[3] Rawan Alharbi, Mariam Tolba, Lucia C. Petito, Josiah Hester, and Nabil Alshurafa. 2019. To Mask or Not to Mask? Balancing Privacy with Visual Confirmation Utility in Activity-Oriented Wearable Cameras. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 72 (Sept. 2019), 29 pages. https://doi.org/10.1145/3351230

[4] Jonas Auda, Uwe Gruenefeld, Sarah Faltaous, Sven Mayer, and Stefan Schneegass. 2023. A Scoping Survey on Cross-reality Systems. *ACM Comput. Surv.* 56, 4, Article 83 (Oct. 2023), 38 pages. https://doi.org/10.1145/3616536

[5] Andrew Besmer and Heather Lipford. 2009. Tagged photos: concerns, perceptions, and protections. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI EA '09)*. Association for Computing Machinery, New York, NY, USA, 4585–4590. https://doi.org/10.1145/1520340.1520704

[6] Virginia Braun and Victoria Clarke. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47. https://doi.org/10.1002/capr.12360

[7] Tamara Denning, Zakariya Dehlawi, and Tadayoshi Kohno. 2014. In situ with bystanders of augmented reality glasses: perspectives on recording and privacy-mediating technologies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 2377–2386. https://doi.org/10.1145/2556288.2557352

[8] R. I. M. Dunbar. 1993. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* 16, 4 (1993), 681–694. https://doi.org/10.1017/S0140525X00032325

[9] Hajer Fradi, Volker Eiselein, Ivo Keller, Jean-Luc Dugelay, and Thomas Sikora. 2013. Crowd context-dependent privacy protection filters. In *2013 18th International Conference on Digital Signal Processing (DSP)*. 1–6. https://doi.org/10.1109/ICDSP.2013.6622808

[10] Hilda Hadan, Derrick M. Wang, Lennart E. Nacke, and Leah Zhang-Kennedy. 2024. Privacy in Immersive Extended Reality: Exploring User Perceptions, Concerns, and Coping Strategies. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 784, 24 pages. https://doi.org/10.1145/3613904.3642104

[11] Abdenour Hadid, Jean-Luc Dugelay, and Matti Pietikäinen. 2011. On the use of dynamic features in face biometrics: recent advances and challenges. *Signal, Image and Video Processing* 5 (2011), 495–506. https://doi.org/10.1007/s11760-011-0247-3

[12] Abdenour Hadid and Matti Pietikainen. 2005. An Experimental Investigation about the Integration of Facial Dynamics in Video-Based Face Recognition. *ELCVIA Electronic Letters on Computer Vision and Image Analysis* 5, 1 (Aug. 2005), 1–13. https://doi.org/10.5565/rev/elcvia.80

[13] Håkon Hukkelås, Rudolf Mester, and Frank Lindseth. 2019. DeepPrivacy: A Generative Adversarial Network for Face Anonymization. In *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Daniela Ushizima, Sek Chai, Shinjiro Sueda, Xin Lin, Aidong Lu, Daniel Thalmann, Chaoli Wang, and Panpan Xu (Eds.). Springer International Publishing, Cham, 565–578. https://doi.org/10.1007/978-3-030-33720-9_44

[14] Håkon Hukkelås and Frank Lindseth. 2023. DeepPrivacy2: Towards Realistic Full-Body Anonymization. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. 1329–1338. https://doi.org/10.1109/WACV56688.2023.00138

[15] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. 2015. Face/Off: Preventing Privacy Leakage From Photos in Social Networks. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (Denver, Colorado, USA) *(CCS '15)*. Association for Computing Machinery, New York, NY, USA, 781–792. https://doi.org/10.1145/2810103.2813603

[16] Mohamed Khamis, Habiba Farzand, Marija Mumm, and Karola Marky. 2022. DeepFakes for Privacy: Investigating the Effectiveness of State-of-the-Art Privacy-Enhancing Face Obfuscation Methods. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (Frascati, Rome, Italy) *(AVI '22)*. Association for Computing Machinery, New York, NY, USA, Article 21, 5 pages. https://doi.org/10.1145/3531073.3531125

[17] Mohamed Khamis, Rebecca Panskus, Habiba Farzand, Marija Mumm, Shaun Macdonald, and Karola Marky. 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 23rd International Conference on Mobile and Ubiquitous Multimedia* (Stockholm, Sweden) *(MUM '24)*. Association for Computing Machinery, New York, NY, USA, 300–312. https://doi.org/10.1145/3701571.3701602

[18] Jan Kietzmann, Linda W. Lee, Ian P. McCarthy, and Tim C. Kietzmann. 2020. Deepfakes: Trick or treat? *Business Horizons* 63, 2 (2020), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006 ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING.

[19] Barbara Knappmeyer, Ian M Thornton, and Heinrich H Bülthoff. 2003. The use of facial motion and facial form during the processing of identity. *Vision Research* 43, 18 (2003), 1921–1936. https://doi.org/10.1016/S0042-6989(03)00236-0

[20] Barbara Knight and Alan Johnston. 1997. The Role of Movement in Face Recognition. *Visual Cognition* 4, 3 (1997), 265–273. https://doi.org/10.1080/713756764

[21] Pavel Korshunov, Andrea Melle, Jean-Luc Dugelay, and Touradj Ebrahimi. 2013. Framework for objective evaluation of privacy filters. In *Applications of Digital Image Processing XXXVI*, Andrew G. Tescher (Ed.), Vol. 8856. International Society for Optics and Photonics, SPIE, 88560T. https://doi.org/10.1117/12.2027040

[22] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. 2017. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 105–114. https://doi.org/10.1109/CVPR.2017.19

[23] Yifang Li, Nishant Vishwamitra, Bart P. Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Blur vs. Block: Investigating the Effectiveness of Privacy-Enhancing Obfuscation for Images. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 1343–1351. https://doi.org/10.1109/CVPRW.2017.176

[24] Yifang Li, Nishant Vishwamitra, Bart P. Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Effectiveness and Users' Experience of Obfuscation as a Privacy-Enhancing Technology for Sharing Photos. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 67 (Dec. 2017), 24 pages. https://doi.org/10.1145/3134702

[25] Chi Liu, Tianqing Zhu, Jun Zhang, and Wanlei Zhou. 2022. Privacy Intelligence: A Survey on Image Privacy in Online Social Networks. *ACM Comput. Surv.* 55, 8, Article 161 (Dec. 2022), 35 pages. https://doi.org/10.1145/3547299

[26] Naresh K. Malhotra, Sung S. Kim, and James Agarwal. 2004. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research* 15, 4 (2004), 336–355. https://doi.org/10.1287/isre.1040.0032

[27] Stanislav Mamonov and Raquel Benbunan-Fich. 2018. The impact of information security threat awareness on privacy-protective behaviors. *Computers in Human Behavior* 83 (2018), 32–44. https://doi.org/10.1016/j.chb.2018.01.028

[28] Richard McPherson, Reza Shokri, and Vitaly Shmatikov. 2016. Defeating Image Obfuscation with Deep Learning. arXiv:1609.00408 [cs.CR] https://arxiv.org/abs/1609.00408

[29] Helen Nissenbaum. 2009. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, Redwood City. https://doi.org/doi:10.1515/9780804772891

[30] Alice J. O'Toole, Dana A. Roark, and Hervé Abdi. 2002. Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences* 6, 6 (2002), 261–266. https://doi.org/10.1016/S1364-6613(02)01908-3

[31] Sang-Min Park and Young-Gab Kim. 2022. A Metaverse: Taxonomy, Components, Applications, and Open Challenges. *IEEE Access* 10 (2022), 4209–4251. https://doi.org/10.1109/ACCESS.2021.3140175

[32] Sarah Prange, Sarah Delgado Rodriguez, Lukas Mecke, and Florian Alt. 2022. "I saw your partner naked": Exploring Privacy Challenges During Video-based Online Meetings. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia* (Lisbon, Portugal) *(MUM '22)*. Association for Computing Machinery, New York, NY, USA, 71–82. https://doi.org/10.1145/3568444.3568468

[33] Franziska Roesner and Tadayoshi Kohno. 2021. Security and Privacy for Augmented Reality: Our 10-Year Retrospective. In *1st International Workshop on Security for XR and XR for Security (VR4Sec)*. https://par.nsf.gov/servlets/purl/10312790

[34] Albrecht Schmidt. 2010. Ubiquitous Computing: Are We There Yet? *Computer* 43, 2 (2010), 95–97. https://doi.org/10.1109/MC.2010.54

[35] Jiayu Shu, Rui Zheng, and Pan Hui. 2018. Cardea: context-aware visual privacy protection for photo taking and sharing. In *Proceedings of the 9th ACM Multimedia Systems Conference* (Amsterdam, Netherlands) *(MMSys '18)*. Association for Computing Machinery, New York, NY, USA, 304–315. https://doi.org/10.1145/3204949.3204973

[36] Samarth Singhal, Carman Neustaedter, Thecla Schiphorst, Anthony Tang, Abhisekh Patra, and Rui Pan. 2016. You are Being Watched: Bystanders' Perspective on the Use of Camera Devices in Public Spaces. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 3197–3203. https://doi.org/10.1145/2851581.2892522

[37] Wouter M. P. Steijn and Anton Vedder. 2015. Privacy under Construction: A Developmental Perspective on Privacy Perception. *Science, Technology, & Human Values* 40, 4 (2015), 615–637. https://doi.org/10.1177/0162243915571167

[38] Jose M. Such, Joel Porter, Sören Preibusch, and Adam Joinson. 2017. Photo Privacy Conflicts in Social Media: A Large-scale Empirical Study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3821–3832. https://doi.org/10.1145/3025453.3025668

[39] Yasuhiro Tanaka, Akihisa Kodate, Yu Ichifuji, and Noboru Sonehara. 2015. Relationship between Willingness to Share Photos and Preferred Level of Photo Blurring for Privacy Protection. In *Proceedings of the ASE BigData & SocialInformatics 2015*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2818869.2818931

[40] Murat Taskiran, Nihan Kahraman, and Cigdem Eroglu Erdem. 2020. Face recognition: Past, present and future (a review). *Digital Signal Processing* 106 (2020), 102809. https://doi.org/10.1016/j.dsp.2020.102809

[41] Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, Aythami Morales, and Javier Ortega-Garcia. 2020. Deepfakes and beyond: A Survey of face manipulation and fake detection. *Information Fusion* 64 (2020), 131–148. https://doi.org/10.1016/j.inffus.2020.06.014

[42] Tomasz Walczyna and Zbigniew Piotrowski. 2023. Quick Overview of Face Swap Deep Fakes. *Applied Sciences* 13, 11 (2023). https://doi.org/10.3390/app13116711

[43] Saima Waseem, Syed Abdul Rahman Syed Abu Bakar, Bilal Ashfaq Ahmed, Zaid Omar, Taiseer Abdalla Elfadil Eisa, and Mhassen Elnour Elneel Dalam. 2023. DeepFake on Face and Expression Swap: A Review. *IEEE Access* 11 (2023), 117865–117906. https://doi.org/10.1109/ACCESS.2023.3324403

[44] Leslie Wöhler, Satoshi Ikehata, and Kiyoharu Aizawa. 2024. Investigating the Perception of Facial Anonymization Techniques in 360° Videos. *ACM Trans. Appl. Percept.* 21, 4, Article 16 (Nov. 2024), 17 pages. https://doi.org/10.1145/3695254

[45] Anran Xu, Shitao Fang, Huan Yang, Simo Hosio, and Koji Yatani. 2024. Examining Human Perception of Generative Content Replacement in Image Privacy Protection. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 777, 16 pages. https://doi.org/10.1145/3613904.3642103

[46] Kaiyu Yang, Jacqueline H. Yau, Li Fei-Fei, Jia Deng, and Olga Russakovsky. 2022. A Study of Face Obfuscation in ImageNet. In *Proceedings of the 39th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 162)*, Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (Eds.). PMLR, 25313–25330. https://proceedings.mlr.press/v162/yang22q.html

[47] Seounmi Youn and Kimberly Hall. 2008. Gender and Online Privacy among Teens: Risk Perception, Privacy Concerns, and Protection Behaviors. *CyberPsychology & Behavior* 11, 6 (2008), 763–765. https://doi.org/10.1089/cpb.2007.0240

[48] Xiaoyi Yu, Kenta Chinomi, Takashi Koshimizu, Naoko Nitta, Yoshimichi Ito, and Noboru Babaguchi. 2008. Privacy protecting visual processing for secure video surveillance. In *2008 15th IEEE International Conference on Image Processing*. 1672–1675. https://doi.org/10.1109/ICIP.2008.4712094

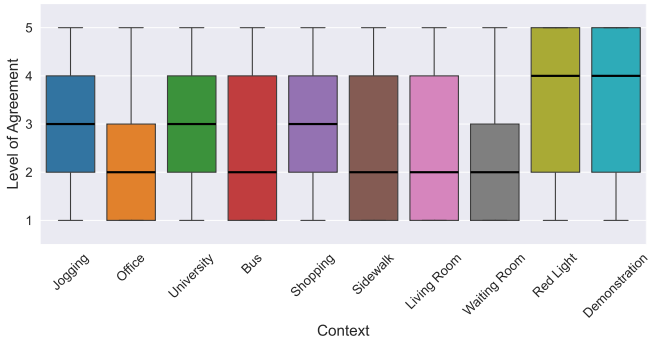
[49] Lin Yuan, Joël Theytaz, and Touradj Ebrahimi. 2017. Context-Dependent Privacy-Aware Photo Sharing Based on Machine Learning. In *ICT Systems Security and Privacy Protection*, Sabrina De Capitani di Vimercati and Fabio Martinelli (Eds.). Springer International Publishing, Cham, 93–107.

A Appendix

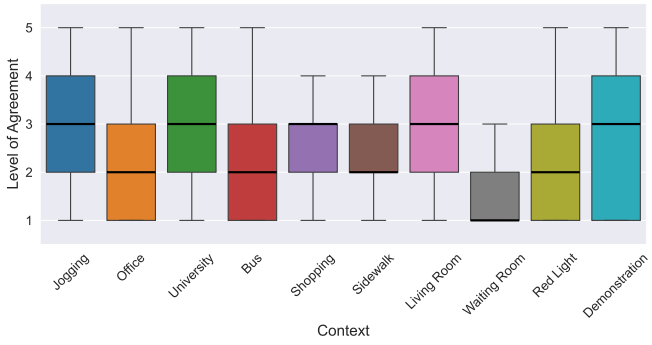
This appendix provides the questionnaires and additional statistical results for the online and laboratory studies.

Table 1: Online Study: Context-Related Questions

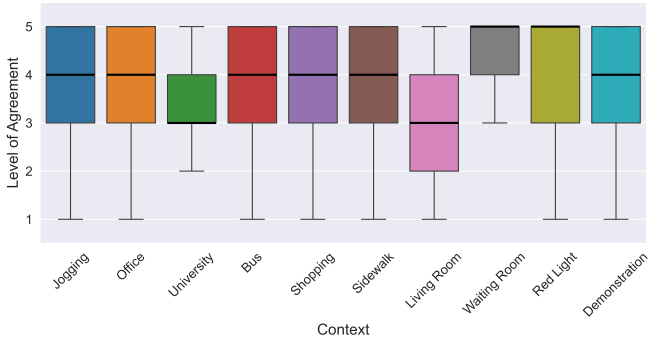
ID	Item	Item Type
Q1	How likely is it to be recorded unintentionally in such a context?	5-point Likert scales from ("1-extremely unlikely" to "5-extremely likely")
Q2	How comfortable are you about being recorded unintentionally in such a context?	5-point Likert scales from ("1-very uncomfortable" to "5-very comfortable")
Q3	I would like to be anonymized and made unrecognizable in the recorded images or video if I am unintentionally recorded in this context.	5-point Likert ("1-strongly disagree" to "5-strongly agree")
Q4	Please indicate how sensitive you consider the information being collected about you when being recorded in such a context. Sensitive information is any information that could cause serious mental, physical, or financial harm if it's lost, misused, or shared without permission.	5-point Likert scales from ("1-not sensitive" to "5-very sensitive")
Q5	Why are you (not) comfortable being recorded in such a context?	Open-ended question
Q6	What would you do if you were recorded in such a context?	Open-ended question



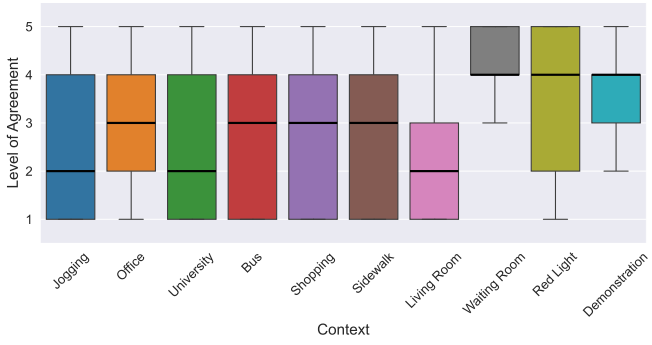
(a) Recording Likelihood



(b) Comfort with Recording

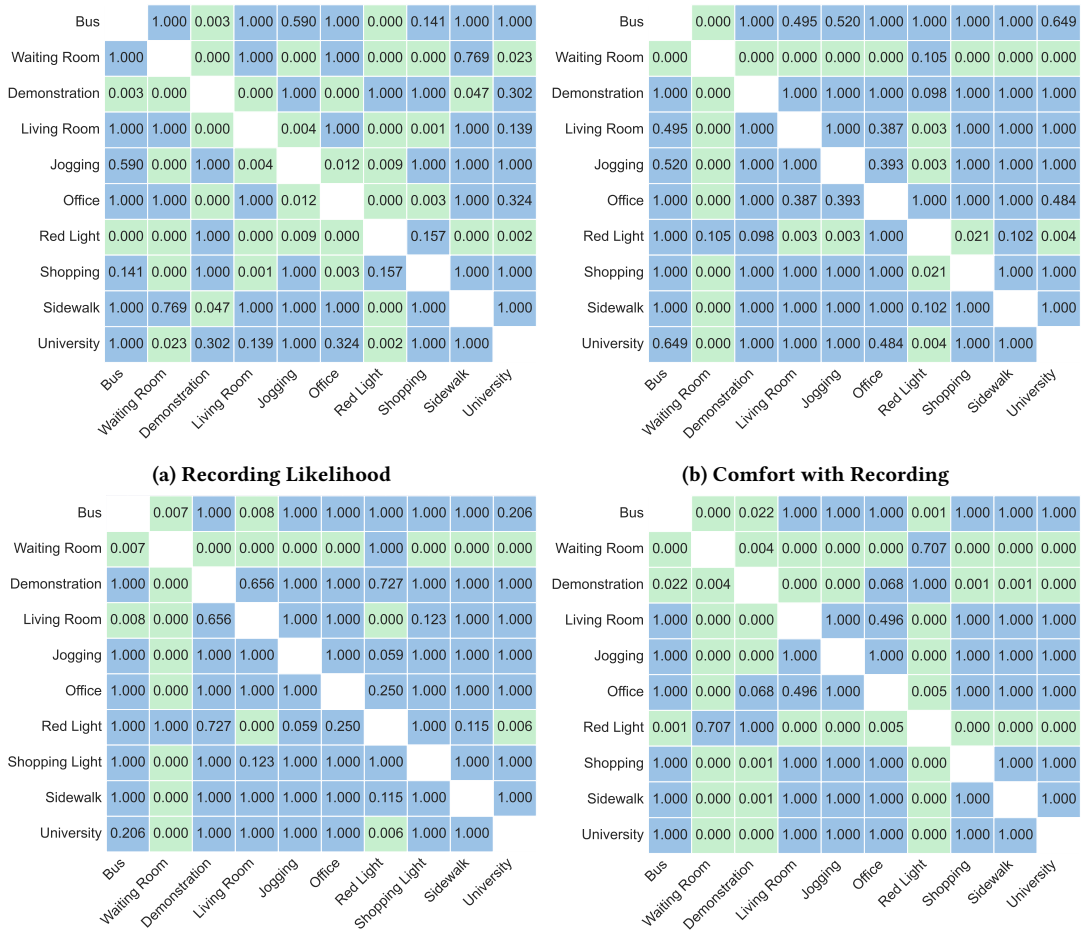


(c) Need for Obfuscation



(d) Sensitivity of Information

Figure 9: Boxplots visualizing the ratings for the different Likert items regarding the sensitivity of being recorded in the context on a 5-Point Likert Scale.



(a) Recording Likelihood **(b) Comfort with Recording**

(c) Need for Anonymization **(d) Sensitivity of Information**

Figure 10: Online Study: P-values (green, < 0.05; blue, >= 0.05) of the Wilcoxon signed ranks tests (Bonferroni correction) for all comparisons for the (a) recording likelihood, (b) comfort with recording, (c) need for anonymization, and (d) sensitivity of disclosed information.

Table 2: Lab Study: Context-Related Questions

ID	Item	Item Type
Q1	How likely is it to be recorded unintentionally in such a context?	7-point Likert scale from 1 (extremely unlikely) to 7 (extremely likely)
Q2	How comfortable are you about being recorded unintentionally in such a context?	7-point Likert scale from 1 (very uncomfortable) to 7 (very comfortable)
Q3	I would like to be anonymized and made unrecognizable in the recorded images or video if I am unintentionally recorded in this context.	7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree)
Q4	Please indicate how sensitive you consider the information being collected about you when being recorded in such a context. Sensitive information is any information that could cause serious mental, physical, or financial harm if it is lost, misused, or shared without permission.	7-point Likert scale from 1 (not sensitive) to 7 (very sensitive)
Q5	Why are you or are you not comfortable being recorded in such a context?	Open-ended question

Table 3: Lab Study: Self-Anonymization-Related Questions

ID	Item	Item Type
Q6	I recognize myself very well in the video material.	7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree)
Q7	To what extent do you find the video with your face edit satisfying?	7-point Likert scale from 1 (very unsatisfying) to 7 (very satisfying)
Q8	I feel my face is well anonymized by this method in this context.	7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree)
Q9	This anonymization method is suitable for this context.	7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree)
Q10	Who would recognize you?	1 (nobody), 2 (myself), 3 (close family), 4 (close friends), 5 (acquaintances), or 6 (strangers)
Q11	I feel uncomfortable with the edited face in the video.	7-point Likert scale from 1 (strongly disagree) to 7 (strongly agree)

Table 4: Lab Study: General and Others-Anonymization-Related Questions

ID	Item	Item Type
Q12	Rank the three anonymization methods based on your preferences for your own anonymization.	Ranking from 1 to 3
Q13	Why did you rank the methods as you did?	Open-ended question
Q7–Q9	For videos showing another person, participants again answered Q7 to Q9 concerning satisfaction, perceived anonymization strength, and contextual suitability.	Same item types as Q7 to Q9
Q14	In videos of myself, I think the anonymization method should be more …	7-point scale from 1 (enjoyable) to 7 (strong)
Q15	In videos of others, I think the anonymization method should be more …	7-point scale from 1 (very enjoyable) to 7 (very strong)
Q16	Why do you prefer strong or enjoyable anonymization for yourself or someone else?	Open-ended question