Impact of Privacy Protection Methods of Lifelogs on Remembered Memories

Passant Elagroudy
German Research Centre for Artificial Intelligence (DFKI)
Kaiserslautern, Germany
LMU Munich
Munich, Germany
passant.elagroudy@gmail.com

Mohamed Khamis
University of Glasgow
Glasgow, United Kingdom
mohamed.khamis@glasgow.ac.uk

Florian Mathis
University of Glasgow
Glasgow, United Kingdom
florian.mathis@glasgow.ac.uk

Diana Irmscher
LMU Munich
Munich, Germany
d.irmscher@campus.lmu.de

Ekta Sood
University of Stuttgart
Stuttgart, Germany
ekta.sood@vis.uni-stuttgart.de

Andreas Bulling
University of Stuttgart
Stuttgart, Germany
andreas.bulling@vis.uni-stuttgart.de

Albrecht Schmidt
LMU Munich
Munich, Germany
albrecht.schmidt@ifi.lmu.de

Figure 1: Original photo of some participants (left) vs. obfuscated version (right) where bodies are blurred (gaussian blur, radius = 40 px). Participants' written consent was obtained to disseminate the photos as part of academic publications.

ABSTRACT

Lifelogging is traditionally used for memory augmentation. However, recent research shows that users' trust in the completeness and accuracy of lifelogs might skew their memories. Privacy-protection alterations such as body blurring and content deletion are commonly applied to photos to circumvent capturing sensitive information. However, their impact on how users remember memories remain unclear. To this end, we conduct a white-hat memory attack and report on an iterative experiment (N=21) to compare the impact of viewing 1) unaltered lifelogs, 2) blurred lifelogs, and 3) a subset of the lifelogs after deleting private ones, on confidently remembering memories. Findings indicate that all the privacy methods impact memories' quality similarly and that users tend to change their answers in recognition more than recall scenarios. Results also show that users have high confidence in their remembered content

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581565

across all privacy methods. Our work raises awareness about the mindful designing of technological interventions.

CCS CONCEPTS

• Security and privacy → Human and societal aspects of security and privacy; **• Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

privacy, filters, obfuscation, lifelogging, blurring, memory augmentation, memory implantation, memory reformation, recall, recognition

ACM Reference Format:
Passant Elagroudy, Mohamed Khamis, Florian Mathis, Diana Irmscher, Ekta Sood, Andreas Bulling, and Albrecht Schmidt. 2023. Impact of Privacy Protection Methods of Lifelogs on Remembered Memories. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3544548.3581565

1 INTRODUCTION

Human memory is naturally prone to reformation over time. Reformation here refers to the decay of memories commonly called forgetting or the memory implantation resulting from suggestability and misattribution [47]. Existing Human-Computer Interaction (HCI) literature traditionally focused on solutions to counter forgetting and augment the human memory (e.g. [40]). Lifelogging is traditionally used in the context of memory augmentation (e.g. [8, 23, 49]). Benefits include: supporting emotional growth, reflection and enhanced reasoning about past experiences, providing motivational cues to future actions, and most importantly supporting the recall of memories. However, the impact of using visually-altered lifelogs on reforming memories intentionally or accidentally remains under-explored.

Research on using technology to prompt memory reformations is currently gaining growing interest [10]. The reformations could be intentional design decisions to alter the users' memories or accidental byproducts of the system design [9]. Examples of deliberate reformations is the disposition of digital content post breakups [25, 26, 46] or maliciously implanting false memories about competitor brands using television advertisements [52]. Although lifelogs are meant as tools for memory augmentation, they can contribute to accidental memory reformations. For example, in one study, eye witnesses removed true details from their statements because they were not captured in their lifelogs from a chest-mounted camera [2]. Similarly, participants did not objectively interpret lifelogs and reported misinformation given by police officers that contradicted cues in their lifelogs [29]. Such reformations could have grave consequences and hinder one of the main motives of lifelogging, that is memory augmentation. However, they remain under-explored in the HCI literature (e.g. [13, 15]).

We hypothesize that such accidental reformations are attributed to users trusting the completeness and accuracy of the digital archives over their natural recall. Thus, the aim of our work here is to do a white-hat memory attack [9] to better understand the impact of using common privacy protection methods of lifelogs on the confidence of lifeloggers while remembering memories, whether skewed or real. Like recent prior work (e.g. [12, 15, 33]), we investigate the impact within the context of *environmental lifelogging*, where wall-mounted cameras capture third-person views of the lifeloggers and bystanders. We chose this context for increased relevance as it mimics surveillance scenarios. Additionally, photos from such infrastructure cameras are clearer than body-worn lifelogs and have higher utility [11], thus better suited for our target. Due to the exposing nature of lifelogs capturing sensitive information, the privacy of the bystanders and lifeloggers is commonly protected in literature and industry using either blurring or deletion of the sensitive content [15, 27, 36, 37, 54]. Thus, our work investigates the impact of both methods on skewing memories.

We report on a between-subject experiment *simulating environmental lifelogging in an artificial lab event* (n=21). The experiment was repeated twice in which participants first took part in an eventful interaction session and then returned after 4-5 days to recall memories when viewing (1) 20 unaltered photos, (2) obfuscated versions of the 20 photos where persons are blurred, and (3) five out of the 20 original photos after deleting private ones. Our experiments are largely inspired by the methodology of Elagroudy et al.'s work [15] and Li et al.'s work [37]. Our work complements [37] by validating the user preferences for the obfuscation methods in the context of lifelogs as opposed to generic photo sharing. Additionally, it significantly complements [15] by: (1) increasing the sample size to test the significance of the reported effects, (2) differentiating recognition and recall contexts while evaluating the privacy-protection method (**PPM**), (3) investigating the impact of confidence in the remembered memories on the memory alterations, and (4) investigating the users' perceived utility of the PPM on enhancing their memories. The term "remember" here is a lay term involving two types of retrieving information; *recognition* and *recall* [16]. We specifically focus on *four* research questions:

RQ1: How does the PPM impact the tendency to change memory narratives?

RQ2: How does the PPM impact the quality of memory narratives?

RQ3: How does the PPM impact the confidence about memory narratives?

RQ4: How helpful are the PPMs as memory prostheses?

We found no evidence of significant difference between the privacy protection methods on the quality of remembered memories. However, users tend to significantly change their remembered memories when exposed to an intervention in recognition situations rather than recall situations. It also shows that users are generally highly confident about their remembered memories when exposed to an intervention (mean = 4, on a 5-Likert value scale; 1=strongly disagree, 5= strongly agree). However, we found no difference in confidence between the three privacy protection methods. Additionally, we learnt that users perceive the intervention photos as unhelpful although our data proves that they enhanced the quality of their remembered memories. These results indicate the relative safety of blurring and deletion as privacy-protection methods that do not hinder the memory utility of lifelogs. Our work is also relevant in the context of media usage from widely-adopted surveillance cameras, which are estimated to be around 1 billion world wide [38]. The described methodology could also be used to investigate the effect of PPMs on regular intentionally-captured personal photos. We envision utilizing our work in: (1) nudging system designers to consider the impact of their presentation decisions on accidental memory reformations, and (2) highlighting a design opportunity to use altered lifelogs to increase users' confidence about skewed memories in deliberate memory alteration scenarios.

2 RELATED WORK

We build on previous work in 1) lifelogs as memory prosthetics and 2) privacy in lifelogs and how it can be protected.

2.1 Lifelogs as Memory Prostheses

Memory cues are stimuli that help individuals retrieve a certain memory. Tulving [57] explained how humans remember information using *the synergistic ecphory* theory where "preconscious process in which retrieval cues are brought in contact with stored information causing parts of that stored information to be reactivated". Cues could be visual (e.g., images), auditory (e.g., sound or speech), locations or mood [17]. Human memory is cue-driven by recognizing such stimuli rather than freely recalling them [51]. Gouveia and Karapanos [18] further suggested, based on Tulving's theory, that cues are of equal strength. However, more cues activate more parts of the episodic memory.

Images promote more detail-rich recall compared to other types of data [18, 30] as they contain rich contextual information [32]. Thus, pictorial lifelogs were extensively researched as a means to augment human memory (e.g., [4, 14, 21, 33, 48, 49]). Pictorial lifelogs could come from wearable cameras capturing first-person perspective excluding the lifelogger or from infrastructure cameras capturing a third-person perspective including the lifelogger. Within We refer to the second type as *environmental lifelogging*.

A large number of works reported on the benefits of existing lifelogging systems in augmenting the human cognition, specifically in the context of memory augmentation (see [8, 23, 49] for examples). Benefits included: supporting emotional growth; reflection and enhanced reasoning about past experiences; providing motivational cues to future actions (e.g., going to the gym); and supporting the recall of memories whether by providing cues to incidents (e.g., checking whether a friend joined a trip) or by repetitively reviewing key incidents to netter remember them.

2.2 Privacy in Lifelogs

The continuous capture of pictorial lifelogs poses significant challenge for the protection of the privacy of bystanders. Privacy infringements could happen through human consumption (others seeing private or uncomfortable content) or via computer vision attacks. This work focuses on the first one. Examples of approaches to protect the privacy include: physically marking objects that should not be captured [43], automated activation of personalized capturing policies in specific contexts [3, 53, 55], automatic or manual deletion of content upon detecting certain cues [31], and obfuscating parts of the photos [31, 36, 37].

We focus in this paper on *two* types of privacy protection: 1) deletion and 2) the obfuscation of photos given that these are most commonly used in literature and practice. Obfuscation is "the production of noise modeled on an existing signal in order to make a collection of data more ambiguous, confusing, harder to exploit, more difficult to act on, and therefore less valuable" [7]. Prior research studied obfuscation of sensitive elements in photos including screens [31], objects [24]) and people [37, 56]. We study the obfuscation of people as they are among the most salient memory cues in photos [32]. Previous work showed that face and body are among the photo parts that are perceived to be sensitive [34]. Thus, obfuscating individuals is done by obfuscating the *face*, where only the head is distorted (e.g., in Google street view), or obfuscating the *body*, where the body and face are distorted [37]. Prior work also showed that cropping bystanders reduces privacy risks while preserving the rest of the image content [56] at the expense of aesthetics. Li et al. evaluated multiple obfuscation methods for face and body. They recommended inpainting individuals or replacing them with avatars because these obfuscation methods provided a good trade-off between effective protection of privacy against computer vision attacks and a good viewer experience [37]. However, they also reported that blurring was among the preferred techniques, conforming with its extensive usage in research and practice [5, 28, 36]. Gross et al. showed that blurring and pixelating of human faces could extensively expose features hindering privacy, or extensively eliminate features hindering the utility of videos [20]. Therefore, we employ body obfuscation in our study as Li et al. found that body obfuscation is more effective against human recognition than face obfuscation only [37].

2.3 Research gap

Sensitive photo elements that violate privacy are often the most valuable memory cues [11]. Although previous works investigated the impact of obfuscation using filters on privacy protection and user experience [24, 31, 36, 37], the impact of obfuscation on the viewer's recall of memories remains under-explored (e.g. [15]). Closing this gap is crucial because recalling memories is one of the main motivations behind lifelogging. Thus in contrast to previous work, we study how the privacy-protection methods (obfuscation and deletion) impact the viewer's remembering of memories.

3 METHODOLOGY

We used a between-subject design to investigate one independent variable, *the privacy-protection method (PPM)*, with the following three conditions and measured their impact on information recall, recognition, confidence about the remembered information, and perceived helpfulness of lifelogs in aiding remembering.

- **C1 (*baseline*)** Participants viewed 20 original photos from the environmental lifelog.
- **C2 (*obfuscation*)** Participants viewed an obfuscated version of (C1), where all persons were blurred (see Figure 1). We used *body blurring* because of the positive results in prior work [5, 28, 36], and because it was the most favored in the first session.
- **C3 (*deletion*)** Participants viewed only a subset of five original photos from (C1), mimicking deletion for privacy protection. We equally sampled them from C1 (*baseline*) dataset across time.

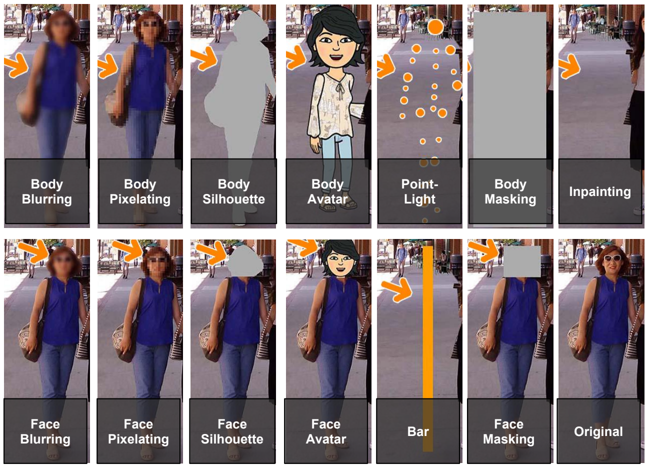

Figure 2: The figure is courtesy of Li et al. [37]. They studied the impact of 13 obfuscation methods on privacy protection and user experience when viewing the above photos. We used the same figures in the context of lifelogging to study the impact of obfuscation on memory recall. In session 1 of our experiment, participants preferred body blurring the most because of its weaker impact on the aesthetics of the lifelogs.

3.1 Experimental Structure

The experiment consisted of two sessions, both of which were conducted in our lab. The *first session* (Session 1) was to create an environmental lifelog of all participants in a common controlled event to evaluate their recall and recognition. We resorted to a synthetic event rather than participants' personal events for ethical reasons as we did not know the scope of the potential memory alterations. The *second session* (Session 2) involved the same participants, and was to compare the impact of the privacy-protection methods on the participants' remembering quality of the events from Session 1, their confidence about them, as well as the perceived utility of the aiding lifelogs.

We conducted two iterations of the same experiment. The rationale was to ensure having a feasible number of participants per session such that: (1) they do not have prior knowledge of each other, and (2) the difficulty of remembering new details and contexts is feasible and balanced. The study design followed the local institutes' research ethics guidelines.

3.1.1 Session 1: Building the Lifelog in a Controlled Event (Group Session). The session lasted for approximately 90 minutes. All participants were invited to the session. We positioned two cameras in the room to record photos from different angles to create an environmental lifelog in third-person view. Participants were informed about the recording, their consent and demographic data was collected, and lifelogging was introduced briefly.

In the first part of session 1, we presented the 13 obfuscation methods [1] studied by Li et al. [37]to the participants (see Figure 2), and asked them for feedback about their most favorite obfuscation method in the context of lifelogging. This was followed by a short open discussion with the participants in the form of a focus group about each technique to understand their rating rationale.

In the second part of session 1, participants were randomly split equally into three teams, each randomly assigned to one of the PPM conditions, to play a board game. In the game, the player's goal is to move their playing pieces to a safe zone based on dice score. We modified the game rules by asking participants to swap seats whenever they score certain dice values. This was done to make the recorded lifelogs more dynamic and to reduce potential bias and confusion from having most of the lifelogs seemingly static when examined in Session 2.

3.1.2 Session 2: Evaluating the Impact on Memory (Individual Session). The second session lasted for approximately 90 minutes. The participants were invited again individually to view lifelogging photos that were taken in session 1. Session 2 took place four to five days after session 1 to ensure a realistic decay of information in the memory [32, 58]. Using a within-subjects design would have led to biases – the information that participants gather in one condition could have influenced their responses in another. To reduce potential bias due to learning effects, we therefore opted for a between-subjects design.

We used a custom-made experimental questionnaire to measure the effects. Each participant answered the questionnaire twice: (1) before viewing the memory cues (pre-questionnaire) and (2) after viewing the memory cues, i.e., photos of the respective condition (post-questionnaire). This was done to account for prior knowledge of the answers and to identify any improvements resulting from having seen the memory cues. Participants were allowed to navigate through the photos as long as they wanted. They were also allowed to improve their answers to the pre-questionnaire when filling the post-questionnaire.

3.2 Experimental Material

We present here the curation strategy for the photos presented in each PPM condition. Additionally, we present the structure of the evaluation questionnaire.

3.2.1 Selection Criteria for Stimuli. In the first session, we collected over 450 photos. We used fixed temporal sampling to select the presented photos (memory cues). We sampled at five-minute intervals for the introduction and obfuscation methods. However, we reduced the interval to three minutes during the game part as it lasted for a shorter period of time (about 20 minutes). Each participant appeared at least once in their experimental lifelog dataset.

Twenty photos were sampled equally from the 450 for the C1 (*baseline*) condition. The selected twenty photos were blurred for the C2 (*obfuscation*) condition. Five photos were sampled equally from the twenty photos to generate the dataset of the C3 (*deletion*) condition.

3.2.2 Questionnaire Structure. Similar to Elagroudy et. al [15], we designed a questionnaire that had 30 questions about details that happened in Session 1. The questionnaire included personal questions (e.g., *"Sophia made very good contributions during the discussion part. Do you remember her field of study?"*), procedural questions (e.g., *"At the start of the study there was one big table in the middle of the room. During the game, we divided this one into multiple smaller*

[1] We obtained the consent of Li et al. to use their figures

tables to distinguish between the different groups. Do you remember the overall number of tables?") and questions about the game (e.g., *"Do you remember which player in your group got a six first? If you are able to recall his/her name please answer with the name. If you can't remember it correctly, try to describe him/her as specific as possible."*).

We mostly maintained the same formulation of questions across both studies while customizing it with details from the session. Example is "X made very good contributions during the discussion part. Do you remember their field of study?", where X was replaced by the name of the participant who spoke the most in the session. To double check the quality and type of the final questions, one researcher designed the questions and labelled their type (*recognition* or *recall*), then another researcher reviewed the questions and independently labelled them as well. The inter-rater agreement was 83%. The labelling conflicts were resolved by a third researcher.

We extended the questionnaire and asked participants to estimate their confidence about their answers ("*I am confident about my answer*) as well as the helpfulness of the photos in aiding recall ("*I found the photos helpful to answer the questions*") during the *post-questionnaire* using 5-point Likert scales (1=Strongly disagree;5=Strongly agree).

3.3 Participants and Recruitment

We had a total of 21 participants (4 females and 17 males). In the first study, we invited 12 participants (3 females and 9 males). In the second study, we invited 9 participants (1 female and 8 males). Each experimental condition had a total of 7 participants from both iterations. The recruitment was done via university mailing list to the students of a computer science faculty and a Facebook social media group for expats (in the city of the study). We ensured to the best of our knowledge that participants do not have prior knowledge of each other. Participants age was between 20 and 32 years old (mean=24.2 years, SD=3.72). We collected demographics such as gender and highest educational degree, and background knowledge such as attitudes towards privacy, and prior knowledge about lifelogging. However, we did not need them later in the analysis. Thus, we skip reporting them here. Participants were compensated with an e-shop voucher (20€). There was a draw on a second voucher to encourage participants to do their best performance in Session 2.

3.4 Evaluation and Analysis

We used six metrics to evaluate the impact of the PPMs on memories. Three metrics are generated through qualitative labelling by the researchers (*Question type labels*, *Changing answers labels* [15], and *Correctness labels*[15]), one metric is automatically calculated via a weighted formula of the assigned correctness labels (*Recall Correctness Score (RCS)*[15]), two metrics are self-reported by the participants in the questionnaire (*Confidence* and *Helpfulness scale*).

Question type labels Two researchers labelled the questions to indicate if the answer is directly available in the photos (*recognition* questions) or it is not presented in the photos (*recall* questions). The questionnaire had 70% *recall* questions (21 questions) and 30% *recognition* questions (9 questions) to mimic information losses during temporal gaps in lifelogs.

Changing answers labels Similar to prior work [15], an answer is considered *changed* in the *post-questionnaire* if the participants provided additional information or edited the answer they had provided in the *pre-questionnaire*. The label is independent from the correctness of the new answer. We refer to questions with changed answers as *changed or reviewed questions*.

Correctness labels Similar to prior work [15, 32], we labelled each answer with one of the following labels: *correct* when the answer is precise, *semi-correct* when some elements of the answer are correct (e.g. describing the gender of a person instead of providing his name), and *wrong* when the answer is irrelevant. We labelled the answers in the *pre-questionnaire* and *post-questionnaire* to gauge the impact of the experimental conditions. The labels were added then discussed by two researchers to ensure internal validity.

Recall Correctness Score (RCS) Following prior work [15], we calculated the RCS to indicate the overall accuracy of recalled memories within a condition. We used the following weights for the correctness labels: *correct*= 2, *semi-correct*= 1, and *wrong*= 0. RCS is a metric corresponding to the summation of weighted correctness labels. Higher values denote better recall quality.

Confidence scale Participants had to indicate their confidence in their response during the *post-questionnaire* on a 5-point Likert scale.

Helpfulness scale Participants had to indicate in the *post-questionnaire* if they found the photos helpful in answering the questions on a 5-point Likert scale.

4 RESULTS

Within this section, we refer to the *type of remembering*, i.e recognition or recall questions as **TR**. All the data used in the research questions was normally distributed and of equal variance, successfully meeting the assumptions of the required tests. The normality was checked using the Kolmogorov-Smirnov test of normality and the variance was checked using Levene's test.

Our null hypothesis H_0 is that *there is no impact from the PPM (and/or the TR when applicable) on the dependant variable.* Proving this null hypothesis would imply designers can use any of the visualizations without fearing their impact on the memories. As frequentist inference allow us to only reject null hypothesis, we also extend our analysis using Bayesian factor analysis[2] to understand the likelihood of this null hypothesis occuring. In the model, we employed a standard non-informative prior where the probabilities are distributed equally. We use Bayesian ANOVA [1, 39, 44] or Bayesian paired sample t-test [1, 39, 45] when appropriate.

4.1 Selected Obfuscation Method: Body Blurring

We asked the participants to select their favorite five obfuscation methods (from the thirteen investigated by Li et al. [35]) in descending order. We counted the frequency (N) of selecting a technique in each of the top three positions (P). Afterwards we calculated a

[2]Ordinal magnitude of evidence ascendingly as used here: no evidence < weak < substantial < strong < very strong < decisive. Some resources group them to weak, moderate and strong.

weighted score for each technique: $Score = 3 * N_{P=1} + 2 * N_{P=2} + 1 * N_{P=3}$

The results showed that *body blurring* was clearly favored (Score= 15), 2) face blurring, body silhouette, body point-light, and body inpainting (Score = 9), 3) face pixelating, face masking, body pixelating, body masking, and body bar (Score= 3), and lastly 4) face silhouette and face avatar (Score= 0) indicating they were never selected in the top three methods. Based on that, we decided to choose body blurring to obfuscate the lifelogs from session 1.

4.2 RQ1: How does the PPM Impact the Tendency to Change Memory Narratives?

A paired-samples t-test was conducted to compare the number of changed answers in all *recognition* and *recall* questions independent of the PPM. The data was normalized[3] to account for the unbalanced number of questions in the experiment. Results show that participants significantly changed their answers more in *recognition* ($M = 0.365, SD = 0.186$) questions more than *recall* ($M = 0.254, SD = 0.115$) questions, ($t(20) = -2.711, p < .05$).

A one-way between-subjects ANOVA was conducted to compare the effect of PPM on the amount of reviewed questions in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). There were no statistically significant differences between the conditions ($F(2, 18) = 1.059, p = .369$).

We also investigated if the results are impacted by the TR. Thus, we separated the number of reviewed questions by question type. We conducted a one-way between-subjects ANOVA two times for *recognition* and *recall* separately to compare the effect of the PPM on the number of reviewed questions in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). There were also no statistically significant differences between the conditions in *recognition* ($F(2, 18) = 0.796, p = .467$) nor in *recall* questions ($F(2, 18) = 0.595, p = .562$).

The Bayesian two-way ANOVA using both PPM and TR as independent variables supplements those results. It shows a strong evidence that the TR affects the number of reviewed answers ($BF_{10} = 13.637$) and a weak evidence that the PPM does not affect the amount of reviewed answers ($BF_{10} = 0.366$). However, there is a weak evidence of an interaction effect between PPM and the TR ($BF_{10} = 1.539$).

Takeaway message: This indicates that participants tended to change their answers and reviewed questions 11% more in *recognition* compared to *recall* questions. However, there is no evidence of one PPM leading significantly to more changes than the others.

4.3 RQ2: How does the PPM Impact the Quality of Memory Narratives?

A one-way between-subjects ANOVA was conducted to compare the effect of PPM on the changes in RCS in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). Higher RCS indicates better quality of memories. We used the difference between the RCS after seeing the photos and before (RCS changes) to exclude remembered content before experiencing the PPM. There were no statistically significant differences between the conditions on RCS changes ($F(2, 18) = 1.121, p = .348$).

We also investigated if the results were impacted by the TR. Thus, we separated the RCS changes by question type. We conducted a one-way between-subjects ANOVA two times for *recognition* and *recall* separately to compare the effect of the PPM on RCS changes in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). There were also no statistically significant differences between the conditions in *recognition* questions ($F(2, 18) = 1.452, p = .26$) nor in *recall* questions ($F(2, 18) = 0.062, p = .94$).

The following Bayesian two-way ANOVA using both PPM and TR as independent variables strengthens those findings. It shows weak evidence that the neither the PPM ($BF_{10} = 0.386$) nor the TR ($BF_{10} = 0.644$) affect the changes in RCS. It also shows a strong evidence of no interaction effect between the PPM and the TR on RCS changes ($BF_{10} = 0.095$).

We also checked if any of the PPMs lead to more answers within specific correctness labels, namely *correct*, *semi-correct*, and *wrong*. This would help us identify if the PPM is better suited for memory augmentation, degradation, or implantation. Thus, we conducted one-way between-subjects ANOVA three times, one for each correctness label to compare the effect of the PPM on the number of questions in each label. The number of questions were normalized based on the number of reviewed questions per participant. There were no significant differences between the PPMs in generating *correct* ($F(2, 18) = 0.273, p = 0.765$), *semi-correct* ($F(2, 18) = 0.871, p = .436$), and *wrong* ($F(2, 18) = 0.546, p = .589$) answers.

We followed up with a Bayesian ANOVA test for each correctness label to clarify the results of the frequentist analysis. It also showed substantial evidence that the PPM did not affect the frequency of *correct*answers ($BF_{10} = 0.307$). Similarly, there was weak evidence that the PPM did not affect the frequency of *semi-correct* ($BF_{10} = 0.434$) and *wrong* ($BF_{10} = 0.36$) answers.

Takeaway message: This implies that given the current data there is no preferable PPM between obfuscation and deletion when it comes to cuing participants to remember better or worse, regardless the questions' type whether it is *recognition* or *recall* .

4.4 RQ3: How does the PPM Impact the Confidence about Memory Narratives?

The mean ratings for the participants' confidence in their remembered answers (from 5-point Likert items) was: C1 (*baseline*)= 4.429, C2 (*obfuscation*)= 4.286, and C3 (*deletion*)= 3.714. The median was 4 in the three conditions. Such high ratings indicate that participants were generally confident about their remembered memories.

A one-way between-subjects ANOVA was conducted to compare the effect of PPM on the confidence of the participants about their remembered memories in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). There were no statistically significant differences between the conditions ($F(2, 18) = 1.432, p = .265$).

We also investigated if the results are impacted by the TR. Thus, we separated the confidence likert-values by question type. We conducted a one-way between-subjects ANOVA two times for *recognition* and *recall* separately to compare the effect PPM on the confidence in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3

[3]Normalization done by calculating a percentage of the reviewed questions from the total number of questions in a category (recognition= 9 questions, recall= 21 questions)

(*deletion*). There were also no statistically significant differences between the conditions in *recognition* questions ($F(2, 18) = 1.465, p = .257$) nor in *recall* questions ($F(2, 18) = 0.659, p = .53$).

The two-way Bayesian ANOVA with PPM and TR as independent variables explains the above results showing a weak evidence that neither the PPM ($BF_{10} = 0.752$) nor the TR ($BF_{10} = 0.344$) on the confidence rating. There is also strong evidence that there is no interaction effect between the PPM and TR affecting the confidence rating ($BF_{10} = 0.073$).

Takeaway message: This implies that all the three PPMs are more likely to have similar capabilities in triggering the participants' confidence about their remembered memories, regardless the questions' type whether it is *recognition* or *recall* .

4.5 RQ4: How Helpful are the PPMs as Memory Prostheses?

We divide the results here into two parts: (1) *objective* efficiency, which refers enhancements in remembering (higher RCS) after seeing the photos, and (2) *subjective* efficiency, which refers to the participants' perception of how much the photos helped them to remember better.

4.5.1 Objective Efficiency of the PPMs. In contrast to Section 4.3, we wanted to check if intervening with photos is generally beneficial in enhancing remembering regardless the type of intervention.

Three paired-samples t-tests were conducted to compare the RCS before and after showing the participants the photos for each of the three PPM conditions. Results show that participants remembered significantly better after seeing the photos compared to before being exposed to the intervention in the three conditions: C1 (*baseline*),C2 (*obfuscation*), and C3 (*deletion*). Table 1 summarizes the means, the standard deviations and the tests' parameters.

We also applied three Bayesian paired-sample t-tests to correspond to the frequentist version. Unlike the frequentist version, there is only weak evidence that participants answer better after seeing the photos in C3 (*deletion*)($BF_{10} = 2.13$). However, there is substantial evidence they remembered better in C2 (*obfuscation*)($BF_{10} = 7.935$), and strong evidence that participants answer better after seeing the photos in C1 (*baseline*)($BF_{10} = 17.469$) . However, as we discussed in Section 4.3, there was no significant difference between the PPM conditions in better promoting higher quality memories.

4.5.2 Subjective Efficiency of PPM. The mean ratings for the helpfulness of the photos (from 5-point Likert items) was: C1 (*baseline*)= 2.786 (median=3), C2 (*obfuscation*)= 2.571 (median=2), and C3 (*deletion*)=2.143 (median=2). Such low ratings indicate that participants generally did not perceive the utility of the photos in aiding them to remember.

A one-way between-subjects ANOVA was conducted to compare the effect of PPM on how helpful the participants perceived the presented photos in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). There were no statistically significant differences between the conditions impacting the helpfulness score ($F(2, 18) = 0.73, p = 0.496$).

We also investigated if the results are impacted by the TR. Thus, we separated the helpfulness likert-values by question type. We conducted a one-way between-subjects ANOVA two times for *recognition* and *recall* separately to compare the effect PPM on the confidence in the three conditions: C1 (*baseline*), C2 (*obfuscation*), and C3 (*deletion*). There were also no statistically significant differences between the conditions in *recognition* questions ($F(2, 18) = 0.711, p = .505$) nor in *recall* questions ($F(2, 18) = 0.813, p = .46$).

The two-way Bayesian ANOVA with PPM and TR as independent variables contradicted these results as it showed a very strong evidence that the TR impacts the subjective helpfulness score ($BF_{10} = 94.012$). Participants found photos more helpful in *recognition* ($M = 3.429, SD = 1.326$) rather than *recall* ($M = 2.048, SD = 0.865$). Conforming with frequentist analysis, the test showed a weak evidence that the PPM does not impact the subjective helpfulness score ($BF_{10} = 0.356$). It also showed a strong evidence of an interaction effect between the PPM and the TR on the subjective helpfulness score ($BF_{10} = 11.322$).

Takeaway message: These results imply that seeing photos led to tangible memory enhancements for the participants. However, the impact of the quantity and quality of the photos remains inconclusive. They also include that while participants did not perceive the photos as useful tools to enhance remembering, they found them more helpful in *recognition* rather than *recall* questions.

5 DISCUSSION

We reflect below on the lessons learnt and limitations of our work. We show that body blurring is a favourable obfuscation method by the users. We also show the potential of using deletion as an obfuscation method as it was not associated with detectable negative effects on remembering nor on user satisfaction.

5.1 Users Prefer Body Blurring as a Privacy-Protection Method

Our results for the most likable obfuscation techniques in the context of environmental lifelogging mirrored those of Li et al. obtained in the context of sharing photos [37]. The participants clearly preferred *body blurring* to other face and body obfuscation techniques. They perceived it as a sufficient technique to protect the privacy and preserve the aesthetics although it is known from prior work that it is ineffective in protecting privacy [24, 34, 35, 37]. However, our results contrast Li et al. in that participants negatively perceived the *avatar* technique [41, 42] although Li et al. [37] recommended it as a good trade-off between efficient privacy protection and positive user experience. This difference is likely due to the context in which the obfuscation was applied, where we considered lifelogs rather than general photo sharing.

5.2 Blurring and Deletion have Good Utility for Memory Prostheses

Our results are inclined towards the lack of difference in the tendency of users to change their answers between deletion, blurring and unaltered photos. Additionally, we cannot detect an impact on the quality of the remembered memories when changing the privacy-protection. Elagroudy et al. [15] suggested a trend where participants tended to generate more semi-correct answers using the obfuscation. The contradiction in our results could be attributed

Condition	RCS Before Condition (Mean,SD)	RCS After Condition (Mean,SD)	Test Report
C1 (*baseline*)	($M = 32.571, SD = 9.431$)	($M = 39.286, SD = 8.261$)	($t(6) = 4.822, p < .05$)
C2 (*obfuscation*)	($M = 31.143, SD = 6.04$)	($M = 36.571, SD = 4.036$)	($t(6) = 3.892, p < .05$)
C3 (*deletion*)	($M = 33, SD = 5.354$)	($M = 36.714, SD = 4.99$)	($t(6) = 2.517, p < .05$)

Table 1: Summary of the PPMs' objective efficiency. Participants remembered better after seeing the photos in all conditions.

to our larger sample size, albeit still relatively small. Therefore, we conservatively recommend that designers can select among the protection methods based on criteria beyond their memory utility. We also encourage future investigations to confirm the trend of current evidence.

5.3 Users Underestimate the Lifelog's Utility to Augment their Memories

Our participants did not think that the original, blurred or deleted lifelogs were helpful in enhancing their recognition and recall. However, objective metrics show that they consistently remembered better after seeing the photos in all conditions. The objective result is not surprising as it conforms with a large body of psychology and HCI literature showing the utility of lifelogs as a memory augmentation tool (e.g. [8, 22, 23, 50]). However, the mismatch underpins a serious challenge in convincing users with the utility of niche technologies such as lifelogging.

5.4 Increasing Photos Quality and Quantity Neither Enhances Memory Augmentation nor User Satisfaction

We expected a clear user preference to photos with higher quantity then quality (original photos, blurred photos, then deleted photos). The expected preference is to conform with the lab study results where the users prioritized aesthetics in selecting the PPM. We also expected increased memory enhancements in the same aforementioned order of PPM conditions. This is based on knowledge from prior literature that memory cues are of equal strength, but that more cues activate more parts of the episodic memory [18]. However, we found no evidence of significant difference in the users' perception of how helpful the photos were regardless their alterations. Additionally, the objective measures of memory enhancements shows no difference between the PPMs as well.

The design implication of these results is that deletion might be superior as a privacy-protection method for memory augmentation systems because it saves storage, offers higher protection for private content, without compromising the lifelogs' utility as memory prostheses. Blurring is also a good option that is widely likable by users and designers alike.

5.5 More Cues Do Not Lead to Higher Confidence in Memories

Participants are equally and highly confident about their answers after viewing original and altered photos. This is an interesting effect as it does not account for the possibility of wrongful cuing from missing or wrong details in the altered photos. The potential confusion can be explained using Gregory's visual assumption theory [19]. He theorizes that the visual perception relies on a top-down approach or conceptual-driven processing where we make calculated assumptions to understand what we see based on our expectations, beliefs and prior knowledge. Similarly, Schachter [47] identifies mis-attribution, i.e., the tendency to confuse the source of a memory with another, and the suggestibility, i.e., the tendency to mix false suggestions by others with the original memory as memory sins. Both of which could be triggered by unclear cues.

The design implication of these results is that blurring could be a superior tool for intentional memory reformation (degradation and implantation) as it provides ambiguous cues that could potentially lead to misattribution while maintaining high confidence of the user in the remembered content. Deletion could also work by systematically omitting information. However, we have no evidence that both PPMs (blurring and deletion) are more dangerous than unaltered photos with respect to accidental memory alterations (augmentation, degradation, and implantation).

5.6 Limitations and Future Work

We would like to acknowledge some limitations to our experiments to better contextualize them. First, we opted for an artificial social event to measure the impact of the recordings on the memory for ethical concerns as the risks associated with the experiment (such as potential memory alterations) were not known. Thus, our approach is a trade-off between short synthetic stimuli like word lists and complex semantic real-world events. Nevertheless, it would be interesting in the future to examine whether the findings hold in complex open-ended memory narratives. Second, we acknowledge that using a custom questionnaire with relatively simplistic questions makes it harder to generalize the results. Third, customizing the questions in the two studies could jeopardize the internal validity of the measurements. However, we circumvented this by consistently using the same format in most of the questions. Fourth, we acknowledge the unequal distribution between the *recognition* and the *recall* questions. This is primarily because we were interested to investigate memory reformation scenarios, which is more common in *recall* situations. We also circumvented this in the analysis by normalizing the results. Fifth, it would be interesting in the future to compare the interventions with a baseline of no photos as we cannot make claims at the moment whether the high confidence is because of the photos or because users are inherently confident about their memories. We would also like to acknowledge the gender disparity in our sample. We postulate this is because of the massively skewed distribution in favour of male students in computer-science domains in Germany. Similarly, the city where the experiment was conducted is highly industrial with a focus on technical jobs, frequently occupied by males. Thus, it is possible that the gender distribution in the Facebook group is also skewed. Nevertheless, it is worth noting that females perform better in episodic tasks than males (e.g. [6, 59]). Thus, the current skew towards males is more conservative for generalizing results, despite the relatively small sample size used in the study.

6 CONCLUSION

Privacy concerns of bystanders and lifeloggers are an ongoing challenge to wider adoption of lifelogging. There is rich literature about the impact of obfuscation on privacy protection and user experience. We complement this work by reporting on two two-stage experiments (N=21) that investigate the impact of privacy-protection methods on the remembering quality of the memories, the users' confidence about the remembered content, and the perceived utility of the privacy methods in the context of environmental lifelogging. We examine two popular methods for privacy protection in literature and practice: obfuscation of persons using body blurring and deletion of private photos as opposed to an original set of photos. Our results show that the privacy method does not impact the quality of the remembered memories in recognition and recall situations. Nevertheless, generally showing users relevant photos encourages them to change their answers in recognition rather than recall situations. Thus, showing external evidence like photos could be a tool for memory reformation. Additionally, results show that users are confident about their responses independent of the privacy-protection method. Users are also unaware about the ambient augmentation effect of lifelogs on their memories and tend to underestimate it. Thus, designers can rule out memory effects when selecting between deletion and blurring to protect the privacy of users. This also implies that deletion is a favourable option to save storage space since it had similar effects on recall and recognition quality like unaltered lifelogs. Our work encourages designers to mindfully account for potential accidental memory effects as a byproduct of designing systems. On the other hand, it encourages future investigations into ambient widely-used visualizations for deliberate memory alterations.

ACKNOWLEDGMENTS

This work is partially supported and funded by the following entities: the Amplify project (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 683008), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC 2075; grant agreement 390740016), the European Research Council (ERC; grant agreement 801708), the University of Edinburgh and the University of Glasgow jointly funded PhD studentships, by an EPSRC New Investigator award (EP/V008870/1) and by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which has been funded by the UK EPSRC under grant number EP/S035362/1, and the HumanE AI Network under the European Union's Horizon 2020 ICT programme (grant agreement no. 952026)
.

REFERENCES

[1] 2017. JASP Team. JASP (Version 0.8.2) [Computer software].
[2] Delene Adams, H. Paterson, and H. MacDougall. 2020. Law and (rec)order: Updating memory for criminal events with body-worn cameras. *PLoS ONE* 15 (2020).
[3] Paarijaat Aditya, Rijurekha Sen, Peter Druschel, Seong Joon Oh, Rodrigo Benenson, Mario Fritz, Bernt Schiele, Bobby Bhattacharjee, and Tong Tong Wu. 2016. I-Pic: A Platform for Privacy-Compliant Image Capture. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services* (Singapore, Singapore) *(MobiSys '16)*. ACM, New York, NY, USA, 235–248. https://doi.org/10.1145/2906388.2906412
[4] Emma Berry, Narinder Kapur, Lyndsay Williams, Steve Hodges, Peter Watson, Gavin Smyth, James Srinivasan, Reg Smith, Barbara Wilson, and Ken Wood. 2007. The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: A preliminary report. *Neuropsychological Rehabilitation* 17, 4-5 (2007), 582–601. https://doi.org/10.1080/09602010601029780 arXiv:https://doi.org/10.1080/09602010601029780 PMID: 17676536.
[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)*. ACM, New York, NY, USA, 4585–4590. https://doi.org/10.1145/1520340.1520704
[6] Sylvia Beyer. 1998. Gender differences in self-perception and negative recall biases. *Sex roles* 38, 1 (1998), 103–133.
[7] Finn Brunton and Helen Nissenbaum. 2015. *Obfuscation: A User's Guide for Privacy and Protest.* The MIT Press.
[8] Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding Quantified-selfers' Practices in Collecting and Exploring Personal Data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. ACM, New York, NY, USA, 1143–1152. https://doi.org/10.1145/2556288.2557372
[9] Sarah Clinch, Omar Alghamdi, and Madeleine Steeds. 2019. Technology-induced human memory degradation. In *Creative Speculation on the Negative Effects of HCI Research. Workshop at The ACM CHI Conference on Human Factors in Computing Systems (CHI 2019).*
[10] Sarah Clinch, Cathleen Cortis Mack, Geoff Ward, and Madeleine Steeds. 2021. *Technology-Mediated Memory Impairment.* Springer Nature, United States, 71–124.
[11] S. Clinch, N. Davies, M. Mikusz, P. Metzger, M. Langheinrich, A. Schmidt, and G. Ward. 2016. Collecting Shared Experiences through Lifelogging: Lessons Learned. *IEEE Pervasive Computing* 15, 1 (Jan 2016), 58–67. https://doi.org/10.1109/MPRV.2016.6
[12] Sarah Clinch, Paul Metzger, and Nigel Davies. 2014. Lifelogging for 'Observer' View Memories: An Infrastructure Approach. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (Seattle, Washington) *(UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 1397–1404. https://doi.org/10.1145/2638728.2641721
[13] Tilman Dingler and Niels Henze. 2014. That's the Dog from my Wedding – Algorithms for Memory Shaping. Workshop Paper. Adjunct Proceedings. CHI Workshop on Designing Technology for Major Life Events. Retrieved June 24, 2016 from https://sites.google.com/site/techmajorlifeevents/Dinger.pdf?attredirects=0.
[14] Aiden R. Doherty, Katalin Pauly-Takacs, Niamh Caprani, Cathal Gurrin, Chris J. A. Moulin, Noel E. O'Connor, and Alan F. Smeaton. 2012. Experiences of Aiding Autobiographical Memory Using the SenseCam. *Human–Computer Interaction* 27, 1-2 (2012), 151–174. https://doi.org/10.1080/07370024.2012.656050 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/07370024.2012.656050
[15] Passant ElAgroudy, Mohamed Khamis, Florian Mathis, Diana Irmscher, Andreas Bulling, and Albrecht Schmidt. 2019. Can Privacy-Aware Lifelogs Alter Our Memories?. In *Proceedings of the 37th Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (Glasgow, Scotland UK) *(CHI EA '19)*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3290607.3313052
[16] Gary Gillund and Richard M Shiffrin. 1984. A retrieval model for both recognition and recall. *Psychological review* 91, 1 (1984), 1.
[17] Lorna Goddard, Linda Pring, and Nick Felmingham. 2005. The effects of cue modality on the quality of personal memories retrieved. *Memory* 13, 1 (2005), 79–86. https://doi.org/10.1080/09658210344000594 arXiv:https://doi.org/10.1080/09658210344000594 PMID: 15724909.
[18] Rúben Gouveia and Evangelos Karapanos. 2013. Footprint Tracker: Supporting Diary Studies with Lifelogging. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. ACM, New York, NY, USA, 2921–2930. https://doi.org/10.1145/2470654.2481405
[19] R. L. Gregory. 1963. Distortion of Visual Space as Inappropriate Constancy Scaling. *Nature* 199 (1963), 678–680.
[20] Ralph Gross, Edoardo Airoldi, Bradley Malin, and Latanya Sweeney. 2006. Integrating Utility into Face De-identification. In *Proceedings of the 5th International Conference on Privacy Enhancing Technologies* (Cavtat, Croatia) *(PET'05)*. Springer-Verlag, Berlin, Heidelberg, 227–242.
[21] Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. 2014. LifeLogging: Personal Big Data. *Foundations and Trends® in Information Retrieval* 8, 1 (2014), 1–125. https://doi.org/10.1561/1500000033
[22] Cathal Gurrin, Alan F. Smeaton, Zhengwei Qiu, and Aiden Doherty. 2013. Exploring the Technical Challenges of Large-Scale Lifelogging. In *Proceedings of the 4th International SenseCam & Pervasive Imaging Conference* (San Diego, California,

USA) *(SenseCam '13)*. Association for Computing Machinery, New York, NY, USA, 68–75. https://doi.org/10.1145/2526667.2526678

[23] Morgan Harvey, Marc Langheinrich, and Geoff Ward. 2016. Remembering through lifelogging: A survey of human memory augmentation. *Pervasive and Mobile Computing* 27 (2016), 14 – 26. https://doi.org/10.1016/j.pmcj.2015.12.002

[24] Rakibul Hasan, Eman Hassan, Yifang Li, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. 2018. Viewer Experience of Obscuring Scene Elements in Photos to Enhance Privacy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 47, 13 pages. https://doi.org/10.1145/3173574.3173621

[25] D. Herron, Wendy Moncur, M. Curic, Drazen Grubisic, Olinka Vistica, and E. V. D. Hoven. 2018. Digital Possessions in the Museum of Broken Relationships. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (2018).

[26] D. Herron, Wendy Moncur, and E. V. D. Hoven. 2017. Digital Decoupling and Disentangling: Towards Design for Romantic Break Up. *Proceedings of the 2017 Conference on Designing Interactive Systems* (2017).

[27] Roberto Hoyle, Robert Templeman, Denise Anthony, David Crandall, and Apu Kapadia. 2015. Sensitive Lifelogs: A Privacy Analysis of Photos from Wearable Cameras. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 1645–1648. https://doi.org/10.1145/2702123.2702183

[28] 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)*. ACM, New York, NY, USA, 781–792. https://doi.org/10.1145/2810103.2813603

[29] K. Jones, W. Crozier, and D. Strange. 2017. Believing is Seeing: Biased Viewing of Body-Worn Camera Footage. *Journal of applied research in memory and cognition* 6 (2017), 460–474.

[30] Vaiva Kalnikaite, Abigail Sellen, Steve Whittaker, and David Kirk. 2010. Now Let Me See Where I Was: Understanding How Lifelogs Mediate Memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. ACM, New York, NY, USA, 2045–2054. https://doi.org/10.1145/1753326.1753638

[31] Mohammed Korayem, Robert Templeman, Dennis Chen, David Crandall, and Apu Kapadia. 2016. Enhancing Lifelogging Privacy by Detecting Screens. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 4309–4314. https://doi.org/10.1145/2858036.2858417

[32] Huy Viet Le, Sarah Clinch, Corina Sas, Tilman Dingler, Niels Henze, and Nigel Davies. 2016. Impact of Video Summary Viewing on Episodic Memory Recall: Design Guidelines for Video Summarizations. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 4793–4805. https://doi.org/10.1145/2858036.2858413

[33] Matthew L. Lee and Anind K. Dey. 2008. Lifelogging Memory Appliance for People with Episodic Memory Impairment. In *Proceedings of the 10th International Conference on Ubiquitous Computing* (Seoul, Korea) *(UbiComp '08)*. ACM, New York, NY, USA, 44–53. https://doi.org/10.1145/1409635.1409643

[34] Yifang Li, Wyatt Troutman, Bart P. Knijnenburg, and Kelly Caine. 2018. Human Perceptions of Sensitive Content in Photos. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

[35] Yifang Li, Nishant Vishwamitra, Hongxin Hu, Bart P. Knijnenburg, and Kelly Caine. 2017. Effectiveness and Users' Experience of Face Blurring as a Privacy Protection for Sharing Photos via Online Social Networks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61, 1 (2017), 803–807. https://doi.org/10.1177/1541931213601694 arXiv:https://doi.org/10.1177/1541931213601694

[36] 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

[37] 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

[38] Marcus Lu. 2022. Ranked: The World's Most Surveilled Cities. Visual capitalist. Retrieved December 12, 2022 from https://www.visualcapitalist.com/ranked-the-worlds-most-surveilled-cities/#:~:text=IHS%20Markit%20estimates%20that%20as,billion%20surveillance%20cameras%20installed%20worldwide..

[39] RD Morey and JN Rouder. 2015. BayesFactor (Version 0.9. 11-3)[Computer software].

[40] Pawarat Nontasil and Stephen J Payne. 2019. Emotional Utility and Recall of the Facebook News Feed. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–9.

[41] José Ramón Padilla-López, Alexandros Andre Chaaraoui, Feng Gu, and Francisco Flórez-Revuelta. 2015. Visual privacy by context: proposal and evaluation of a level-based visualisation scheme. *Sensors* 15, 6 (2015), 12959–12982.

[42] Chi-Hyoung Rhee and C LEE. 2013. Cartoon-like avatar generation using facial component matching. *Int. J. of Multimedia and Ubiquitous Engineering* 8, 4 (2013), 69–78.

[43] Franziska Roesner, David Molnar, Alexander Moshchuk, Tadayoshi Kohno, and Helen J. Wang. 2014. World-Driven Access Control for Continuous Sensing. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (Scottsdale, Arizona, USA) *(CCS '14)*. ACM, New York, NY, USA, 1169–1181. https://doi.org/10.1145/2660267.2660319

[44] Jeffrey N Rouder, Richard D Morey, Paul L Speckman, and Jordan M Province. 2012. Default Bayes factors for ANOVA designs. *Journal of mathematical psychology* 56, 5 (2012), 356–374.

[45] Jeffrey N Rouder, Paul L Speckman, Dongchu Sun, Richard D Morey, and Geoffrey Iverson. 2009. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review* 16, 2 (2009), 225–237.

[46] Corina Sas and Steve Whittaker. 2013. *Design for Forgetting: Disposing of Digital Possessions after a Breakup*. Association for Computing Machinery, New York, NY, USA, 1823–1832. https://doi.org/10.1145/2470654.2466241

[47] Daniel L. Schachter. 2002. *The seven sins of memory*. Boston: Houghton Mifflin.

[48] Abigail J. Sellen, Andrew Fogg, Mike Aitken, Steve Hodges, Carsten Rother, and Ken Wood. 2007. Do Life-logging Technologies Support Memory for the Past?: An Experimental Study Using Sensecam. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. ACM, New York, NY, USA, 81–90. https://doi.org/10.1145/1240624.1240636

[49] Abigail J. Sellen and Steve Whittaker. 2010. Beyond Total Capture: A Constructive Critique of Lifelogging. *Commun. ACM* 53, 5 (May 2010), 70–77. https://doi.org/10.1145/1735223.1735243

[50] Abigail J Sellen and Steve Whittaker. 2010. Beyond total capture: a constructive critique of lifelogging. *Commun. ACM* 53, 5 (2010), 70–77.

[51] Roger N. Shepard. 1967. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6, 1 (1967), 156 – 163. https://doi.org/10.1016/S0022-5371(67)80067-7

[52] Susan M. Sherman, Hannah Follows, Alexander B.R. Mushore, Kathleen Hampson-Jones, and Katie Wright-Bevans. 2015. Television advertisements create false memories for competitor brands. *Journal of Applied Research in Memory and Cognition* 4, 1 (2015), 1–7. https://doi.org/10.1016/j.jarmac.2014.06.001

[53] Julian Steil, Marion Koelle, Wilko Heuten, Susanne Boll, and Andreas Bulling. 2018. PrivacEye: Privacy-Preserving First-Person Vision Using Image Features and Eye Movement Analysis. *CoRR* abs/1801.04457 (2018). arXiv:1801.04457 http://arxiv.org/abs/1801.04457

[54] Julian Steil, Marion Koelle, Wilko Heuten, Susanne Boll, and Andreas Bulling. 2019. PrivacEye: Privacy-Preserving Head-Mounted Eye Tracking Using Egocentric Scene Image and Eye Movement Features. In *Proceedings of the 2019 ACM Symposium on Eye Tracking Research and Applications (ETRA'19)*. ACM, New York, NY, USA. https://doi.org/10.1145/3314111.3319913

[55] Robert Templeman, Roberto Hoyle, David Crandall, and Apu Kapadia. 2014. Reactive Security: Responding to Visual Stimuli from Wearable Cameras. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (Seattle, Washington) *(UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 1297–1306. https://doi.org/10.1145/2638728.2641708

[56] Edison Thomaz, Aman Parnami, Jonathan Bidwell, Irfan Essa, and Gregory D. Abowd. 2013. Technological Approaches for Addressing Privacy Concerns when Recognizing Eating Behaviors with Wearable Cameras. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. ACM, New York, NY, USA, 739–748. https://doi.org/10.1145/2493432.2493509

[57] Endel Tulving. 1982. Synergistic ecphory in recall and recognition. *Canadian Journal of Psychology/Revue canadienne de psychologie* 36, 2 (1982), 130.

[58] John T. Wixted and Ebbe B. Ebbesen. 1991. On the Form of Forgetting. *Psychological Science* 2, 6 (1991), 409–415. https://doi.org/10.1111/j.1467-9280.1991.tb00175.x

[59] A Daniel Yarmey. 1993. Adult Age and Gender Differences in Eyewitness Recall in Field Settings 1. *Journal of Applied Social Psychology* 23, 23 (1993), 1921–1932.