See Me If You Can: A Multi-Layer Protocol for Bystander Privacy with Consent-Based Restoration

Yahya Khawaja*
Computer Science
Lahore University of Management Sciences (LUMS)
Lahore, Pakistan
26100137@lums.edu.pk

Shirin Rehman*
Computer Science
Lahore University of Management Sciences (LUMS)
Lahore, Pakistan
26100164@lums.edu.pk

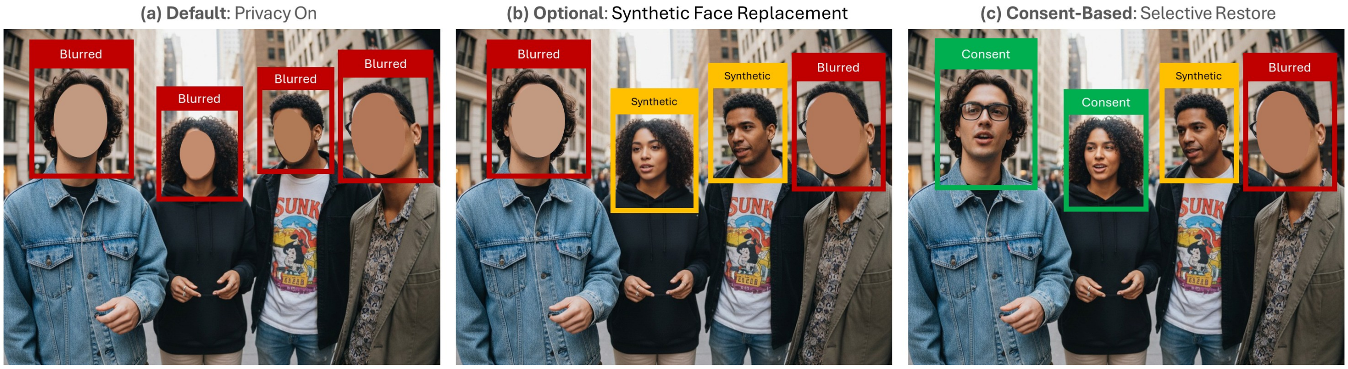
Alexander Ponticello
CISPA Helmholtz Center for Information Security
Saarbrücken, Germany
alexander.ponticello@cispa.de

Divyanshu Bhardwaj
CISPA Helmholtz Center for Information Security
Saarbrücken, Germany
divyanshu.bhardwaj@cispa.de

Katharina Krombholz
CISPA - Helmholtz Center for Information Security
Saarbrücken, Germany
krombholz@cispa.de

Muhammad Hamad Alizai
Computer Science
Lahore University of Management Sciences (LUMS)
Lahore, Pakistan
hamad.alizai@lums.edu.pk

Naveed Anwar Bhatti
Computer Science
Lahore University of Management Sciences (LUMS)
Lahore, Punjab, Pakistan
naveed.bhatti@lums.edu.pk



(a) Default: Privacy On

(b) Optional: Synthetic Face Replacement

(c) Consent-Based: Selective Restore

Figure 1: First-person capture with our three-tier protocol: (a) all faces are blurred on-device by default; (b) selected faces are replaced with identity-neutral synthetic faces; (c) after per-person consent, only consenting faces are restored while others remain protected.

Abstract

The growing popularity of wearable camera glasses raises pressing concerns about bystanders being recorded without their consent. Most existing privacy-enhancing technologies (PETs) rely on opt-out models that place the burden of privacy protection on bystanders. We conducted a qualitative study on wearers' and bystanders' perceptions of opt-in, privacy-by-default approaches for camera glasses. To enable this study, we designed and evaluated an opt-in privacy-by-default protocol. We then conducted semi-structured interviews with camera glass wearers and bystanders ($N = 18$) to examine their perceptions of the protocol. Our findings show that bystanders viewed the opt-in protocol as essential and advocated for even stronger anonymization. Wearers appreciated the protocol's safeguards but found it visually limiting, expressing desire for a context-dependent version that can be enabled in relevant scenarios. Our findings highlight the need for context-aware PETs that provide effective mechanisms for consent negotiation.

*Both authors contributed equally to this research.

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3790394

CCS Concepts

• **Human-centered computing** → **Ubiquitous and mobile computing**; • **Security and privacy**;

Keywords

Camera glasses, Wearables, Opt-in, Privacy-Enhancing Technologies

ACM Reference Format:
Yahya Khawaja, Shirin Rehman, Alexander Ponticello, Divyanshu Bhardwaj, Katharina Krombholz, Muhammad Hamad Alizai, and Naveed Anwar Bhatti. 2026. See Me If You Can: A Multi-Layer Protocol for Bystander Privacy with Consent-Based Restoration. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3772318.3790394

1 Introcution

Camera glasses now resemble conventional eyewear so closely that they often go unnoticed in everyday settings. Early prototypes, like Google Glass [10], stood out with their conspicuous design, making it clear when someone was using them. In contrast, modern products such as those produced by Meta in collaboration with Ray-Ban [18] are styled to blend in seamlessly. This subtlety makes it easier for wearers to record daily interactions without drawing attention, but it also raises pressing concerns about privacy, as bystanders may be filmed without their awareness or explicit consent.

Prior research has documented how bystanders experience unease and discomfort around camera glasses, citing ambiguity about when and how recording occurs [8]. In response, researchers have proposed various privacy-enhancing technologies (PETs), such as automated face blurring and blocking tools [41, 57]. While these tools provide baseline protection, they erase valuable social cues such as expressions [46], diminishing recording value while offering no avenues for content restoration. Other PETs exploring opt-out mechanisms prove impractical because they require bystanders to actively signal their privacy preferences [25, 34]. Existing opt-in mechanisms, such as explicit hand gestures, require real-time negotiations between wearers and bystanders. These interactions can be socially awkward, ambiguous in meaning, and provide no enforceable privacy guarantees. Practical solutions must therefore address the needs of both groups, support consent negotiation, and establish enforceable privacy protections [3, 61]. Opt-in[1] privacy-by-default mechanisms are well-suited to meet these requirements, yet remain under-explored.

We conduct a qualitative user study to examine how wearers and bystanders perceive opt-in, privacy-by-default PETs for wearable camera glasses. In particular, we answer the following research questions:

RQ1 What are the privacy needs of bystanders in the context of privacy-by-default, opt-in mechanisms for wearable camera glasses?

RQ2 What are the usability requirements of wearers in the context of privacy-by-default, opt-in mechanisms for wearable camera glasses?

To address these questions, we required a concrete opt-in, privacy-by-default PET. Existing solutions were either opt-out, lacked enforceable protections, or did not offer consent-based restoration. Since no system satisfied these requirements, we developed a novel opt-in privacy-by-default protocol for camera glasses to enable our subsequent exploration of user perceptions. Our protocol applies automatic face blurring at the point of capture and supports consent-based restoration. It also offers AI-based synthetic face replacement to enhance media quality without compromising bystander privacy. We implement and evaluate our protocol to demonstrate a realizable system rather than a purely speculative design.

Using this protocol, we conduct a qualitative study with N=18 participants (9 wearers and 9 bystanders). Participants interacted with the protocol interface and discussed their perceptions in semi-structured interviews.

Our work makes two key contributions. *First*, we provide an empirical account of how wearers and bystanders experience and negotiate opt-in, privacy-by-default PETs for camera glasses, focusing on their privacy needs, usability requirements, and the social dynamics of consent-based restoration. *Second*, we implement and evaluate a novel three-tier protocol, demonstrating that an opt-in privacy-by-default design is technically viable on wearable-class hardware. To our knowledge, this is the first study to examine both wearer and bystander perceptions of a privacy-by-default protocol for camera glasses that combines default on-device obfuscation with consent-based restoration.

Our findings reveal three key insights: (1) consent-based restoration can effectively mediate privacy needs, but its usage for both wearers and bystanders is highly context-dependent; (2) bystanders strongly supported mandatory opt-in obfuscation mechanisms; and (3) wearers found certain protocol features cumbersome but were receptive to a more context-dependent way of applying and relaxing the protocol.

These findings inform the design of opt-in, privacy-by-default PETs and highlight implications for devices that balance meaningful recording with the rights of those recorded. We also highlight how technology-mediated consent mechanisms intersect with the often implicit social negotiations surrounding privacy.

2 Related Work

Research on privacy for camera glasses spans HCI, computer vision, security, and policy. To situate our contribution, we structure prior work into four areas: usage and adoption, bystanders' perspectives, wearers' perspectives, and privacy-enhancing technologies.

2.1 Usage of Camera Glasses

Camera glasses build on a long history of wearable innovation [8, 30, 48], evolving from early prototypes such as Google Glass and Snapchat Spectacles [10, 47] to recent Meta–Ray-Ban collaborations [18, 31]. A central shift in this trajectory has been design philosophy: whereas early devices were visibly futuristic, newer models are fashionable and often indistinguishable from regular eyewear [3,

[1] *Opt-in* refers to approaches where privacy protection is the default and restoration occurs only with explicit bystander consent. In contrast, *opt-out* refers to approaches where recording proceeds by default and bystanders must actively intervene to restrict it.

10, 18, 31]. Zuidhof et al. argue that this transition is not merely aesthetic but sociotechnical, as these devices mediate relations between wearers and bystanders [61], and their inconspicuousness is central to modern privacy challenges [15].

Although research highlights diverse functional applications of smart glasses [2, 29, 42, 58], adoption is driven largely by social use. Bipat et al. [4] found that wearers primarily value hands-free, first-person video for natural and authentic documentation. These same qualities raise ethical and privacy risks: the discreet form factor obscures recording, leaving bystanders uncertain if they are being captured [15]. Images may reveal sensitive details about locations, activities, or associations without consent [52, 59], and advances in real-time facial recognition make large-scale identification increasingly feasible [7, 34, 52], eroding anonymity in public. Early lifelogging studies showed that in the absence of safeguards, users improvised social protocols, such as turning cameras away or verbally negotiating consent, to manage bystander privacy [16].

2.2 Bystanders' Perspectives

Bystanders have been the subject of several studies regarding camera glasses and other wearable recording devices. Denning et al. [8] established that bystander reactions to these devices are often negative and rooted in the fact that camera glasses could record them with ease and subtlety. A key finding was that bystanders desire to be asked for permission before being recorded. A large-scale survey by Niu et al. [33] demonstrated that a person's role as a bystander does not directly correlate with their level of discomfort. Instead, privacy preferences are dictated by the context of the situation, such as the perceived sensitivity of the activity being recorded.

A study by Koelle et al. [22] on consent mechanisms found that bystanders feel social discomfort when using explicit opt-out gestures toward being recorded, and instead may silently accept privacy infringements. Zhao et al. [60] studied bystanders with visual impairments and found that standard visual privacy indicators, such as LED lights, are inaccessible to this group. Another study on user attitudes towards camera glasses found that bystanders had significantly more negative attitudes towards the glasses compared to smartphones or webcams, particularly in public contexts, and therefore recommended that they only be used for specialized tasks with clear purposes [23].

2.3 Wearers' Perspectives

As wearable recording devices have become increasingly common over the past decade, researchers have explored their usage and broader implications in real-world contexts [23, 43]. Hoyle et al. [16] investigated the use of lifelogging devices and found that users actively avoid sharing images that might compromise bystander privacy. Bipat et al. [4] found that the use of camera glasses is shaped by both personal preferences and prevailing social norms, with wearers actively assessing the contextual appropriateness of a situation before recording.

Bhardwaj and Ponticello et al. [3] further investigated the ethical burden placed on wearers by current camera glasses and found that many wearers feel personally responsible for protecting bystander privacy. In the absence of adequate technical solutions, participants reported developing their own mitigation strategies, such as recording under non-sensitive circumstances only or simply not wearing the glasses under certain situations. The study highlighted the need for a privacy-preserving mechanism that removes this burden from the wearers and advocated for the development of privacy-mediating technologies that do not rely solely on interpersonal trust to manage these tensions.

While prior work has begun to examine how wearers navigate the social and ethical complexities of smart glasses, it highlights a critical gap: the need for systems that relieve wearers of the responsibility of managing bystander privacy.

2.4 Existing Privacy Enhancing Technologies (PETs)

Current camera glasses rely on audio and visual cues such as capture LEDs and clicking sounds to signal recording. These cues are intended as privacy indicators but have proven ineffective. Portnoff et al. [39] showed that fewer than 50% of users noticed LEDs during computer tasks and under 5% during written tasks. Koelle et al. [22] similarly criticized LEDs for poor visibility, clarity, and trustworthiness. Bhardwaj and Ponticello et al. [3] confirmed that both audio and visual indicators fail to reliably notify bystanders. Collectively, these studies show that current indicators fall short in addressing bystander concerns.

PETs attempt to address this gap but remain limited. Systems such as ScreenAvoider [24] protect specific contexts, such as preventing cameras from capturing sensitive computer screens, while others (FaceBlock, SnapMe, SelfFlag, I-Pic [1, 13, 27, 57]) require bystanders to broadcast policies. These assume static preferences, lack restoration if consent is later granted, and rely on an opt-out model where recording occurs by default. This shifts the burden to bystanders, who must install apps, wear devices, or use gestures, creating major usability challenges.

Cloud-based systems such as EgoBlur [41] adopt a privacy-by-default approach by blurring all faces. While protective, this method only permits permanent blurring and depends on external servers, raising concerns about exposing sensitive data. Krombholz et al. [25] evaluated PET concepts including a Privacy App, Privacy Fabric, and Privacy Bracelet, all requiring bystander action. Participants were particularly wary of the Privacy App's reliance on third-party servers, viewing it as a serious privacy risk.

3 Protocol Design

To ground our exploration of **RQ 1** and **RQ 2** in empirical insights, we propose and implement a novel opt-in, privacy-by-default protocol for camera glasses [19]. Our implementation enables our user study by demonstrating the feasibility of such PETs. We outlined the following design goals based on the current limitations of camera glass PETs outlined in Section 1 and Section 2.4:

RQ1 *Privacy by Default:* The protocol must follow an opt-in model that ensures bystander privacy is protected automatically, without requiring any indication of preference or explicit action from the wearer or bystander.

RQ2 *Dual Support:* The protocol must balance the needs of both wearers and bystanders by providing mechanisms that safeguard bystander privacy while providing a practical recording experience for wearers.

RQ3 *Minimize Third-Party Reliance:* The protocol must not rely on sharing recorded media with any third party.

Based on our design goals, we develop an on-device privacy-enhancing obfuscation mechanism where all media recorded by camera glasses will have the faces of bystanders blurred by default. This shifts the responsibility for privacy protection away from both the wearer and the bystander and instead places it on the system itself. We chose blurring as our default obfuscation method due to its effectiveness in providing a baseline level of obfuscation [26, 54] and its ability to be performed on small devices, such as camera glasses. Additionally, our protocol offers two alternative options for users to replace the blurred faces in the recorded media. We first introduce a novel consent-based face restoration mechanism that relies on a Trusted Third Party (TTP) to send consent requests to bystanders. In line with our design goals, no recorded media is shared with the TTP during this process. Prior research highlights that wearers often find it tedious or infeasible to obtain consent in crowded or dynamic environments, where interactions may last only a few seconds [3]. Our approach allows wearers to restore the original media later if consent is granted. Leveraging a TTP provides a practical solution, as the widespread adoption of digital platforms ensures that most individuals can be easily onboarded. Secondly, we include a feature that allows users to use AI-based synthetic face replacement to replace the blurred face with an AI-generated face. Existing research shows that synthetic face replacement can be an effective method for improving video quality while also protecting bystander privacy [12, 55].

3.1 Threat Model

Our protocol is designed for both wearers and bystanders by providing strong privacy and security guarantees for bystanders and suitable alternatives for wearers. While a determined wearer could bypass protections by simply using an unrestricted device, we nonetheless adopt a rigorous threat model that includes adversarial wearers to ensure strong privacy guarantees under well-defined conditions. Side-channel attacks and direct firmware compromise are out of scope. We focus on three primary adversaries:

1. **Malicious Wearer:** The attacker has access to the smart glasses and the companion smartphone. They may jailbreak the phone to attempt to extract raw unblurred frames, or identify bystanders. However, the wearer cannot compromise the smart glasses themselves.
2. **Compromised Cloud Operator:** The TTP handling consent is assumed honest-but-curious, yet vulnerable to external attack or legal coercion. Such an adversary may inspect server memory, storage, or traffic to obtain blurred or unblurred media. Cloud operators colluding with malicious wearers is out of scope and the wearer's companion mobile device does not allow the companion app to share recorded media with the cloud operator.
3. **On-Path Network Attacker:** Monitors, intercepts, or modifies traffic between glasses, phone, and cloud. Their goal is to intercept recorded media or encrypted data and attempt to decrypt it offline.

Based on our design goals and threat model, we implemented our protocol through a three-tiered architecture shown in (Figure 2).

3.2 Tier 1 - On-Device Blurring (Camera Glasses)

Tier 1 operates entirely on the smart glasses and implements the core obfuscation mechanism. For every recorded frame (image or video), a lightweight face detector identifies bounding boxes for each face. For each detected face, three key operations are performed:

1. **Landmark Extraction:** Facial landmarks are extracted using a lightweight model. These will later support expression-preserving AI face replacement.
2. **Embedding Generation:** A compact face embedding [44] is computed to uniquely (and irreversibly) represent each bystander. For videos, we use a tracking-based approach to only compute one embedding for a face stream (same face across frames). This embedding will later be used to identify the bystander for potential consent requests. In the case of a person leaving and re-entering the video, the tracking-based approach will consider this as a new face stream (distinct bystanders). We will discuss how to fix this issue in Section 3.4.
3. **Encryption:** Each person's face region is encrypted using a unique symmetric key (AES 128 bits from the OS's CSPRNG) generated randomly for that individual. Using one key per person enables selective restoration of specific individuals. This same key is also used to encrypt the corresponding face embedding, ensuring consistent and secure encryption across frames. To protect this symmetric key, it is then encrypted using the public key of a TTP. This ensures that only the TTP can decrypt the key and that the wearer cannot access the original faces or embeddings without consent.

Finally, a Gaussian blur is applied to each detected face in the frame 3. Sensitive data like encrypted symmetric keys are stored in the glasses' Trusted Execution Environment (TEE) [40]. The following data is then transmitted to the wearer's phone upon pairing and subsequently deleted from the camera glasses: the blurred media, encrypted facial regions, encrypted face embeddings, unencrypted facial landmarks, and the symmetric key (encrypted with the TTP public key). Facial landmarks remain unencrypted, as they are not typically considered to contain personally identifiable information. Although some research [50] suggests that identification from landmarks may be possible in rare forensic scenarios with small suspect pools; this risk is significantly lower on a larger scale such as ours. As an additional security measure, we add adversarial systematic random noise like shifting key features like the nose or eyes by notable offsets to further obfuscate the data.

A crucial security consideration for Tier 1 computation is the mechanism by which the camera glasses obtain and maintain the TTP's public key. Because the glasses lack native Wi-Fi or cellular connectivity and therefore cannot obtain keys directly, key distribution must occur through the paired mobile device.

The TTP's public key is selected from a pool of valid keys that the glasses obtain by establishing a secure Bluetooth channel with

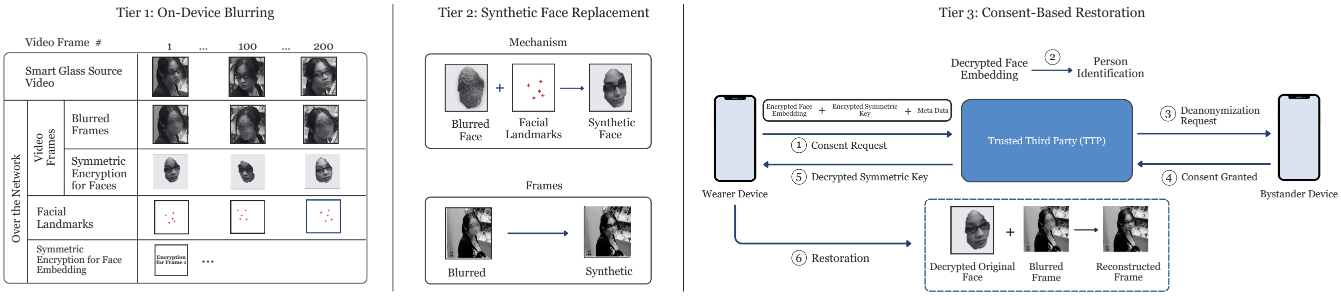


Figure 2: Overview of the three-tier privacy protocol: Tier 1: on-device blurring with encrypted embeddings; Tier 2: optional synthetic face replacement; Tier 3: consent-based restoration via a TTP with steps 1 to 6 marked.

the paired mobile device. The mobile device (which has full network connectivity) retrieves the current list of TTP public keys from a designated Key Management Server (KMS) or the TTP's infrastructure. This list is digitally signed by the camera manufacturer's Root Key. The device then transfers the signed key list to the glasses over the secure Bluetooth channel. The glasses verify the manufacturer's signature using the Root Public Key that is permanently embedded in the firmware. If verification succeeds, the glasses accept and use the new TTP public key for wrapping bystander encryption keys.

This dynamic key-delivery process supports secure key rotation and revocation. To prevent a malicious wearer from evading updates, the camera glass firmware is required to periodically request a fresh signed key list. Since embedded root public keys are intended to be long-lived, rotating them is out of scope, but still possible if the vendor embeds multiple root keys and uses a controlled update to deprecate one and activate another.

3.3 Tier 2 - Synthetic Face Replacement (Companion Phone)

Tier 2 runs on the companion phone and can be used to improve the viewing experience without compromising privacy. It uses the blurred media from Tier 1 along with the unencrypted facial landmarks to replace AI-generated faces for each blurred region. These synthetic replacements preserve natural facial expressions and movements while anonymizing the bystanders' identity. We chose synthetic face replacement to provide a visually consistent alternative to simple blurring or consent requests and support contexts where aesthetic or interpretability concerns matter (e.g., lifelogging, productivity reviews) 3.

3.4 Tier 3 - Consent-Based Restoration (TTP Coordination)

Tier 3 allows the wearer to restore the original faces in the video if bystanders provide their consent via a TTP 3. To give consent, a bystander must be a registered member of the TTP platform. The consent request process is as follows:

(1) The companion phone sends only the encrypted face embedding(s), encrypted symmetric key(s), and metadata (discussed in section 6.4) about the media to the TTP via a secure SSL channel along with an authentication token only known to camera glasses and the paired mobile device, generated upon initial pairing (Figure 2, Tier 3-Step ①). The camera glasses will have added this token to the encrypted data at the point of capture. The TTP authenticates this request by matching the authentication token shared by the mobile device with the token added by the camera glasses (after decrypting it). This ensures only the paired mobile device can make a valid consent request for the associated recording.

(2) The TTP decrypts the symmetric key(s) using its private key, then uses the symmetric key(s) to decrypt the corresponding embedding(s). If there are multiple embeddings (in case of a video), the TTP will compute the similarity between these embeddings, and if it finds two or more embeddings with a high similarity threshold, it will merge their corresponding face streams (face across frames) into one stream and treat them as a single bystander. This will resolve the issue with the same bystander leaving and re-entering the video. The final embedding(s) are then compared with a reference database of embeddings to identify the bystander(s) (Figure 2, Tier 3-Step ②). Crucially, the identification process is governed by a strict, pre-defined similarity threshold to mitigate the risk of misidentification. Only embeddings that exceed this threshold are considered a match and sent a consent request. This threshold should be finetuned to ensure that ambiguous matches are deliberately filtered out and never proceed to the consent stage, thus providing a strong privacy guarantee against inadvertently asking non-matching or misidentified individuals for permission. The specific threshold depends on the similarity matching algorithm/model. For our proof of concept, we used a naive cosine similarity algorithm with a 0.65 threshold based on initial testing on a prototype database of 15 face embeddings.

(3) Upon finding a match, the TTP sends a consent request to the identified bystander(s) through their profiles on the TTP's platform (Figure 2, Tier 3-Step ③). Once this match is identified, the TTP must delete the face embeddings.

In order to give consent bystanders must enroll their face embeddings through an initial verification process when signing up with the TTP. This verification can be carried out through an identity authentication mechanism such as an active liveness check. This authentication mechanism will ensure fake accounts cannot be used to grant consent, as the embeddings will be computed during the active liveness check. The embeddings could also be computed from the existing profile pictures of the bystanders on the TTP platform (e.g., a trusted entity like Meta or Google), but it is crucial to have another robust authentication mechanism for fake profiles in this scenario [21, 45]. It is important to note that the embeddings cannot be computed without the explicit knowledge and informed consent of the person. Additionally, if a bystander has not opted into the system or does not have a registered profile with the TTP platform, they will not receive consent requests. As a result, their face will remain blurred, preserving privacy by default 3.

If a match is found, the bystander is sent a consent request through the TTP platform. The exact medium for the consent request could vary, such as the TTP profile, email, or SMS. The consent request includes metadata about the recording. This metadata may include the identity of the requesting wearer, a timestamp, duration, approximate location, or a personalized message from the user describing the purpose of the request. We do not specify the exact metadata included since it represents a crucial tradeoff between contextual transparency (to support informed consent from the bystander) and minimizing third-party exposure (to uphold design principle three). This tradeoff is discussed in section 6.4 and left open for further research and community debate.

If the bystander grants consent (Figure 2, Tier 3-Step ④), the TTP transmits the decrypted symmetric key securely (e.g., via SSL) to the companion phone (Figure 2, Tier 3-Step ⑤). This companion phone is identified as the same device that made the initial consent request. Note: In the case of a video where the same bystander enters and exits the frame repeatedly, the TTP will send the keys for all face streams corresponding to that bystander. It is important to note that these symmetric key(s) are only valid for the bystander who provided consent and cannot be used to decrypt face regions for other bystanders in the media. The symmetric keys should be deleted by the TTP after successful transmission; regardless, our split-key protocol design ensures the symmetric keys cannot be used maliciously by the TTP. The phone then decrypts the corresponding face regions using these key(s) and merges them into the original media, thus restoring the original face (Figure 2, Tier 3-Step ⑥). Crucially, all decryption and restoration processing occurs locally on the wearer's phone 3.

The number of requests the TTP must handle is a key scalability factor for Tier 3 computation. We make several optimizations to mitigate this overhead: Tiers 1 and 2 operate entirely on the glasses and companion phone, minimizing TTP involvement, and Tier 3 computation occurs only when the wearer requests consent. Additionally, the TTP only processes encrypted keys, embeddings, and metadata instead of raw media, significantly reducing data transfer and network latency overheads. Finally, its computation is limited to lightweight similarity matching, keeping compute costs low.

3.5 Protocol Limitations

Since reliably identifying and obtaining consent for a specific person's audio stream across a range of different scenarios is technically not feasible at this point in time, we do not include audio privacy as part of our obfuscation-restoration mechanism. However, this is a technological limitation rather than an architectural limitation, and our protocol design supports an audio obfuscation-restoration mechanism as the technology progresses.

Additionally, our protocol relies on transmitting unencrypted facial landmarks to the mobile device to support synthetic face replacement. Although landmarks are not typically treated as personally identifiable information and we take steps to reduce associated risks, some potential for re-identification still remains. Future work should address this limitation by exploring landmark-free approaches to synthetic face replacement.

Furthermore, our protocol temporarily stores encrypted face regions to support future consent requests. Although this data is encrypted, it introduces potential risks if, for example, the TTP's private key is compromised or the encryption scheme becomes vulnerable. To limit this exposure, the wearer's companion app automatically deletes all encrypted face regions or embeddings if consent is not requested or granted within a fixed time window. We also acknowledge the possibility of the TTP attempting to maliciously exfiltrate encrypted data from the wearer's phone via the companion app, but these edge cases are blocked by the companion phone's strict permission model, which prevents the app from transferring such data. TTP-Wearer collusion could also present a new threat angle, where either party may break the key-data split but such cases are beyond the scope.

4 Implementation and Evaluation

We implemented our protocol on wearable-class hardware and evaluated performance across blurring accuracy for Tier 1, visual fidelity for Tier 2, and associated system cost for Tier 1. Our goal is to confirm the technical viability of privacy-by-default architectures on resource-constrained devices and provide a concrete, realizable foundation for our subsequent investigation of user perspectives.

4.1 Implementation

The main feasibility concern lies in whether camera glasses can handle the on-device computation described in Tier 1. To assess this, we implement Tier 1 of our protocol on a Raspberry Pi 4 Model B [38] with a Pi Camera module V1.3 [49], serving as a stand-in for the Snapdragon AR1 Gen 1 [40] platform that powers the current Meta Ray-Ban collection [18]. Table 1 shows a comparison of the computational capability of both platforms. The Raspberry Pi has a slightly slower CPU, lacks dedicated machine learning accelerators such as NPUs or tensor cores, and its VideoCore VI GPU is limited to basic video processing and lightweight 3D rendering. The Tier 1 pipeline uses a three-thread model: a read thread loads frames into a queue, a processing thread performs blurring and encryption and passes results to a write queue, and a write thread collects the frames from the write queue and stores them. This prevents processing delays from blocking frame reading.

We implement Tier 2 using conventional face warping methods and MobileFaceSwap (MFS) [56], a mobile-optimized deep learning architecture for face swaps. For each bystander, a suitable synthetic identity (based on skin tone and landmark alignment) is selected from a database of AI-generated faces. Once a synthetic face is chosen, we apply a geometric warping pipeline using affine transformations to align the synthetic face on top of the blurred area and then apply MFS (with the same synthetic face) to reduce visual artifacts such as discontinuities and misaligned features.

Table 1: Hardware specifications: AR1 Gen 1 vs RPi 4 Model B.

Specification	AR1 Gen 1	RPi 4 Model B
CPU	4 cores @ **1.9 GHz**	4 cores @ **1.8 GHz**
Architecture	64-bit	64-bit
GPU	Adreno	VideoCore VI
NPU	**Yes**	**No**
Tensor Accel.	**Yes**	**No**
Memory	LPDDR4	LPDDR4

We implement Tier 3 by simulating the role of the TTP on a remote server equipped with a prototype database of 15 bystanders. Once a bystander provides consent, we execute the following decryption and restoration mechanism through a companion Android application.

4.2 Evaluation

We structure our evaluation across three dimensions: privacy protection, visual utility, and system cost. Privacy protection quantifies the effectiveness of our Tier 1 blurring pipeline using Average Precision and Average Recall, following the MS-COCO protocol [28]. We compare the results of our privacy protection evaluation against EgoBlur, a SOTA face blurring PET [41]. Visual utility measures the perceptual quality of Tier 2 synthetic replacement by using standard perceptual metrics like Fréchet Inception Distance [14] and Structural Similarity Index [51]. Since our synthetic face replacement pipeline operates on blurred faces, we compare our performance against a baseline of applying MFS directly on unblurred faces. System cost captures the latency and energy overhead of Tier 1 on the Raspberry Pi 4 during blurring and encryption compared to a baseline of regular video recording (protocol disabled). We also evaluate the storage overhead for encrypted face packets compared to a baseline average video size. Privacy protection and visual utility evaluations were conducted on a custom curated dataset of videos recorded using the Rayban Meta Stories glasses [31]. The dataset was categorized across a varying number of bystanders, distances, and movement scenarios (see GitHub [20]). System cost evaluations were conducted on live videos (following the same categorization) recorded using the Pi Camera Module in order to avoid underestimating camera I/O costs. Table 2 shows summarized results from all three evaluation dimensions.

Our results demonstrate that opt-in, privacy-by-default smart glasses are practically feasible, even when evaluated on a relatively constrained development platform. The Tier 1 privacy pipeline demonstrates robust efficacy, with the initial blurring stage achieving high precision and recall comparable to existing methods, even within constrained environments. The secondary synthetic replacement stage successfully maintains high visual fidelity, yielding reconstruction quality competitive with SOTA swapping techniques on unblurred faces. While implementing these protection mechanisms introduces moderate increases in processing latency, energy consumption, and storage requirements compared to baseline operations, the computational overheads remain manageable. Furthermore, the protocol ensures efficient resource utilization by imposing no latency and storage burdens, alongside only a minimal increase in energy usage when no bystanders are detected. Additional implementation details and extended quantitative results for each dataset category are available in our GitHub repository [20].

5 Study

We conducted a qualitative study to examine wearer and bystander perceptions of opt-in, privacy-by-default mechanisms for camera glasses. Our protocol enabled this investigation by providing participants with a functional system in order to ground their responses in a realizable system.

The study employed a two-phase design with two participant groups: wearers and bystanders. Wearers were provided with the Meta Ray-Ban Stories [31] during a one-week onboarding period. Since identifying and recruiting natural smart-glass users was not feasible due to the limited adoption of such devices, the onboarding period served as a methodologically grounded alternative that simulated real-world use. It allowed wearer participants to become familiar with the technology and provided a realistic experience akin to that of someone who had recently acquired camera glasses, thereby improving the ecological validity of our study. Following onboarding, wearers participated in an in-person interview session. Bystander participants proceeded directly to the in-person interview, as their role required no prior device exposure.

5.1 Recruitment

We recruited participants from our university community. While we recognize that this population is not representative of all demographics, it is particularly well-suited for examining initial perceptions of opt-in privacy mechanisms in camera glasses, as this demographic represents early technology adopters likely to encounter such systems as they emerge [37].

We employed a non-probability sampling strategy that combined purposive and quota sampling to ensure diversity in our participant pool. For both groups, we began recruitment by posting in the university's private student Facebook group (Figure 3 summarizes the recruitment process), directing interested participants to a screening survey. The screening survey collected demographic data, consent information, and participant preference for being recruited as a wearer or bystander. The survey also contained a detailed description of the tasks expected from both groups. While we accepted participants on a rolling basis, we used quota sampling to ensure gender balance among both wearer and bystander groups [35]. Additionally, to achieve greater diversity in age and educational background, we employed purposive sampling to directly contact university faculty and staff [36].

Neither participant group was informed of the study's purpose until the interview session to reduce potential bias. This was particularly important for the wearer group, as we aimed to avoid influencing their onboarding usage towards a privacy-aware perspective. To mitigate the risk of information spreading between potential participants due to our localized, university-based recruitment, we recruited across departments and asked participants not to share details about the protocol or interview process. We continued recruiting participants in each group until the primary interviewers agreed that no substantial new themes emerged, indicating that data saturation had been reached.

Table 2: Evaluation results across all three dimensions. Arrows (↑ /↓) indicate the better direction. Theoretical ranges are shown in brackets. For III System Cost, Priv. Only specifies the system cost for the core obfuscation-restoration mechanism. Priv. + Syn. includes the additional cost of landmark detection needed to enable the synthetic face replacement feature. No Byst. specifies system costs when no bystander is being recorded.

I. Privacy Protection (Tier 1)			**II. Visual Utility (Tier 2)**			**III. System Cost**				
Metric [Range]	**EgoBlur**	**Ours**	**Metric [Range]**	**Baseline**	**Ours**	**Metric**	**Baseline**	**Priv. Only**	**Priv. + Syn.**	**No Byst.**
AP ↑ [0, 1]	0.935	**0.942**	FID ↓ [≥ 0]	**31.00** ± 13.8	63.70 ± 27.8	Storage (MB)	56.16	69.33	69.33	56.16
AR ↑ [0, 1]	**0.974**	0.953	SSIM ↑ [0, 1]	**0.76** ± 0.07	0.61 ± 0.07	Energy (J)	40.00	67.04	112.05	49.66
			PSNR (dB) ↑ [0, ∞)	**15.87** ± 2.77	12.85 ± 1.95	Latency (s)	9.98	13.69	22.88	10.02

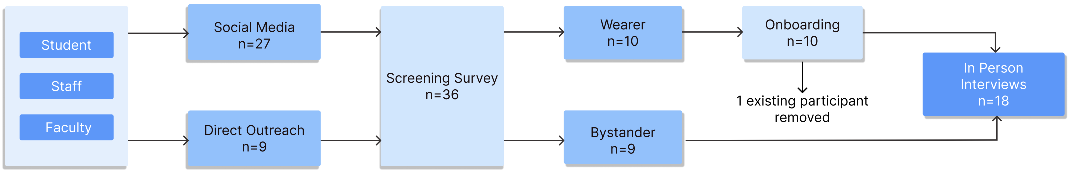


Figure 3: Overview of our recruitment and assignment workflow including number of participants at each stage.

Table 3: Participant Demographics and Details

	Participant	Age	Gender	Education	Usage
Wearers	W1	18–24	M	High school	7 days
	W2	18–24	M	High school	7 days
	W3	18–24	M	High school	11 days
	W4	25–34	F	Graduate	12 days
	W5	25–34	F	Graduate	13 days
	W6	18–24	F	High school	9 days
	W7	18–24	M	High school	20 days
	W8	18–24	F	High school	20 days
	W9	25–34	M	Bachelors	11 days
Bystanders	B1	18–24	M	High school	–
	B2	18–24	F	Bachelors	–
	B3	45–54	M	Graduate	–
	B4	25–34	M	Graduate	–
	B5	25–34	M	Graduate	–
	B6	18–24	M	Graduate	–
	B7	45–54	M	Graduate	–
	B8	18–24	M	High school	–
	B9	18–24	F	High school	–

Note. Education reflects the highest completed degree. If, e.g., participants were currently enrolled in a bachelor's program but had not yet completed it, their education is reported as high school.

The final sample comprised 18 participants (9 Wearers, 9 bystanders; see Table 3 and Figure 3). One additional wearer enrolled but withdrew during onboarding due to personal reasons and contributed no data; they are excluded from all analyses. Ages are reported in ranges to protect participant privacy and comfort.

5.2 Data Collection

Apparatus. We used two pairs of Meta RayBan Stories sunglasses [31] with identical form except for color. The sunglass form factor was deemed appropriate as we anticipated primary use in outdoor settings would align naturally with the local climate and participants' daily activities. Although this form factor may constrain indoor usage, we determined that this limitation did not compromise the study's exploratory objectives. The glasses feature dual 5MP cameras, onboard storage, audio capabilities, and a capture LED for privacy indication. They allow first-person photo and video capture, with media stored locally and triggered via a temple button.

5.2.1 *Phase 1: Onboarding Phase (Wearer Only).*

To ground insights in real-world use, wearer participants were provided with a pair of factory-reset Meta RayBan Stories for a minimum of seven days. While the target duration was one week, some participants retained the glasses for longer due to scheduling conflicts. Wearers were instructed to use the glasses as naturally as possible in their daily lives and document their experiences in daily diary entries which they were encouraged to keep in order to motivate regular use of the glasses. We deliberately crafted instructions to avoid privacy-related themes in order to minimize potential bias. A private WhatsApp chat group with three members (the participant and two lead researchers was created for each wearer participant to facilitate diary submission and provide regular reminders. Participants submitted entries at their convenience and in their preferred format (text message, voice note, etc.) to minimize frustration and fatigue [6]. Participants also received an additional PKR 200 for each diary entry submitted, capped at seven days.

5.2.2 *Phase 2: In-Person Session.*

All participants took part in an in-person session composed of two stages: a protocol demonstration followed by a semi-structured interview. Sessions were conducted by a team of two interviewers.

Protocol Demonstration. Section 4 presented our protocol implementation on Raspberry Pi hardware. However, since the user study required authentic camera glass capture, we used Meta Ray-Ban glasses as an abstraction for the users. Participants recorded short videos using the glasses, with footage processed through our protocol to generate blurred and AI-replaced versions. These processed recordings were then displayed via a custom web application that served as the protocol's user-facing interface (Figure 4). This hands-on interaction grounded subsequent interview responses in personal experience while maintaining experimental consistency across all sessions. For wearers, participants recorded the primary interviewer; for bystanders, the interviewer recorded the participant. These recordings served as the basis for demonstrating the protocol, with the reversed recording direction ensuring each participant experienced the system from their stakeholder perspective.

Following the recording, the session proceeded in parallel: one interviewer presented a visual slide deck explaining the protocol and answering participant questions, while the second interviewer processed the recorded video through the protocol pipeline.

Participants then used a custom web application that served as the protocol's user interface (Figure 4). The interface then guided participants through each protocol tier sequentially:

(1) Participants viewed their recordings with all faces blurred (Tier-1), which represented the privacy-by-default state.
(2) Participants then initiated Tier-2, which displayed the original video with AI-replaced faces.
(3) Participants could also request restoration (Tier-3), which triggered a pop-up detailing the number of bystanders identified and how consent requests would be delivered. For the sake of the demo, the consent response was set to true. Upon confirming, participants could view the original recording.

Across all tiers, participants could pause, playback, and zoom in on the recordings with each interactive session taking approximately 15 minutes. This experience with the interface provided participants with the necessary understanding of the protocol design before proceeding to the interview. This follows established methods in prior research, where hands-on experience with technology grounds subsequent discussion [17].

Interview. The interview was designed to elicit in-depth feedback on the privacy protocol aligned with our research questions. Participants discussed each protocol tier individually, then evaluated the integrated system. For wearers, interviewers focused on usability and comfort; for bystanders, on privacy needs across different scenarios. The interview concluded with questions on participants' overall impressions, allowing them to reflect on perceived strengths, weaknesses, and their specific privacy and usability priorities. Further details on the interview guidelines can be found in Table 4. Each participant received PKR 1000 compensation upon completing the interview.

To reduce potential bias, participants were told that the protocol was developed by an unaffiliated third party. This framing encouraged participants to focus on their genuine perceptions rather than assumptions about the researchers' expectations or involvement.

5.3 Pilot Testing

To validate and refine our study design, we conducted a series of pilot tests. Initial pilots on a volunteer participant and the researchers themselves helped identify and resolve technical issues with the protocol demonstration and address minor flaws in the protocol explanation, such as ambiguous language. The key takeaway from the pilot phase was that participants gave the most valuable insights when they reflected on the protocol from their own perspective, either wearer or bystander, rather than trying to imagine what others might think. While perspective-switching can yield interesting observations, we did not want it to become our primary focus. To support this, we adjusted the wording in the protocol explanation and interview questions to keep the focus on their specific role and minimize perspective-switching during responses.

5.4 Data Processing

All interview audio recordings were transcribed using a GDPR-compliant AI transcription service. The interviewer who conducted the session then manually reviewed the entire transcript against the audio to correct errors. This manual review was necessary to accurately capture participants' code-switching, as all 18 interviews were conducted bilingually in English and Urdu. The researchers, who are proficient in both languages, provided contextually accurate English translations for all Urdu portions of the interview.

5.5 Thematic Analysis

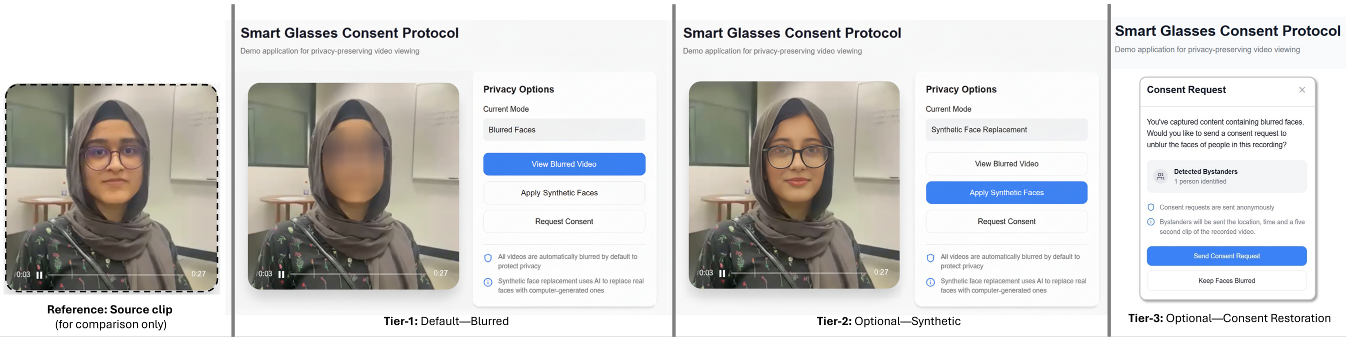
We employed an inductive qualitative analysis approach that follows the thematic analysis method by Braun and Clark [5]. We chose this approach because it has been effective for previous exploratory studies regarding camera glasses [3, 4, 9]. The two primary researchers independently coded the same two interviews using an open coding approach. They then met to compare and merge their codes, producing the initial codebook. Any disagreements were resolved collaboratively, and the codebook was refined accordingly. We determined that this process was sufficient to allow the remaining transcripts to be divided between the two researchers. The finalized codebook is provided in the supplementary material.

5.6 Ethical Considerations

This study was approved by the Institutional Review Board (IRB) at our university. All participants were given informed consent forms prior to the in-person evaluation sessions, including consent for audio recording and transcription. Wearer participants were explicitly informed that they were not required to share any photos or videos taken during their use of the smart glasses. However, some chose to voluntarily include media in their diary entries. They were also given the option to factory reset the glasses before returning them, ensuring that all locally stored data and media would be deleted. Participants were informed that interviews would be audio-recorded and that these recordings would be stored securely, with access restricted to the two lead researchers. The short video captured during the interview for the protocol demonstration was used solely for that purpose and was permanently deleted immediately after the session.

5.7 Study Limitations

Given the limited natural adoption of camera glasses, we provided participants with Meta Ray-Ban glasses during a one-week onboarding phase. This approach offered them a realistic experience similar to that of someone newly acquiring camera glasses, but it also introduced limitations, such as potential artificial behavior resulting from the short exposure period. However, because the focus of



Demo application for privacy-preserving video viewing

Current Mode: Blurred Faces

All videos are automatically blurred by default to protect privacy

Synthetic face replacement uses AI to replace real faces with computer-generated ones

Current Mode: Synthetic Face Replacement

You've captured content containing blurred faces. Would you like to send a consent request to unblur the faces of people in this recording?

Detected Bystanders: 1 person identified

Consent requests are sent anonymously

Bystanders will be sent the location, time and a five second clip of the recorded video.

Figure 4: Three-tier privacy workflow in the demo viewer. *Left*: reference frame (for comparison only, not shown to participants). *Tier-1 (Default)*: participants viewed faces blurred. *Tier-2*: participants toggled to synthetic; identity-neutral faces replace real faces to preserve watchability. *Tier-3*: participants click "Send Consent Request" to trigger a simulated consent request; only consenting faces can be locally restored. All interface elements were participant-controlled.

our study was to gain initial perceptions about a novel technology, this methodological approach is supported by precedents in prior research [3, 37].

All participants were residents of Pakistan. As with any localized study, cultural context may influence findings. Pakistan's unique socio-cultural landscape, including norms around privacy, gender dynamics in public spaces, and technology adoption patterns, may shape participant perspectives differently than populations in other regions. Future work should validate these findings across diverse geographic and cultural contexts.

Our study was designed to capture initial perceptions and immediate usability feedback but does not observe behavioral patterns over time. This strategy is consistent with established methodologies for emerging camera glass technologies [11, 17]. Future work should validate these findings through longitudinal studies using real-world deployments of opt-in systems.

Additionally, our study focused on a single protocol rather than providing participants with multiple alternative designs. Although this approach allowed for hands-on interaction with a functional system, it may have biased participant responses to the specific design choices embedded in our protocol. Therefore, our findings should be understood as grounded in one realizable implementation rather than an exhaustive exploration of the design space for opt-in, privacy-by-default PETs. The limited sample size of our study may also limit the generalizability of the findings to the broader population.

6 Results

We structure our results section to effectively answer our research questions: RQ1 examines bystander privacy needs in opt-in approaches, while RQ2 investigates wearer usability requirements. Our findings reveal that consent mechanisms introduce complex social dynamics, obfuscation effectiveness depends heavily on context, and both stakeholder groups balance competing priorities: bystanders emphasize privacy protection while wearers prioritize usability and recording capability. We use **W1-9** to denote wearer participants and **B1-9** to denote bystander participants. Omissions and edits for readability are denoted with square brackets.

6.1 Dynamics of Consent Requests

The consent mechanism directly addresses both research questions: for wearers (RQ2), it shaped usability perceptions and recording decisions; for bystanders (RQ1), it influenced privacy protection expectations. Participants expressed a nuanced understanding of the consent mechanism, illustrating that the decision to request, grant, or deny consent was shaped by social relationships, contextual appropriateness, and the perceived burden of interaction.

6.1.1 When Wearers Decide to Send Requests. Wearers described a range of considerations that shaped their decision to send a consent request. Rather than treating it as a routine action, they weighed how likely the request would be accepted, who the recipients were, and whether the context justified having the original faces. These judgments reflected both the social dynamics of their relationships and the situational importance of the recording.

Likelihood of Acceptance. Wearers took into consideration two main factors when weighing the probability of having their consent requests accepted:

(1) *Relationship with Recipient.* Wearers were most willing to request consent in situations involving friends, family, or small gatherings, where the request could be understood as part of a shared social activity.

> *[W1]"I would only be sending out requests when I am with friends or family and I have some semblance of hope that they would consent."*

Most wearers strongly objected to sending requests to strangers, who were described as unlikely to accept such requests. W1 questioned why anyone would accept a request from a stranger, believing only *"a very small minority"* would accept. Many wearers also felt that asking for consent from strangers would be uncomfortable or even inappropriate, particularly in the absence of what they thought to be valid justification.

(2) *Ability to Act on Requests.* Wearers also emphasized that they would only send consent requests to people who were familiar with the protocol and capable of responding to it.

W4 specified that they would not send requests to individuals who *"might not be using [the protocol]," "might not know about it", "might be busy,"* or *"might not have their phones with them."* They also believed that *"parents and [older people]"* would likely not be able to give consent because *"they're not using much of these apps."*
Thus for wearers, the willingness to send a request was also influenced by their perception of the recipient's ability to engage with the consent mechanism.

Content of Recording. Additionally, wearers' decisions to send consent requests were shaped by the content of the recordings themselves.

(1) *Significance of Recording.* Wearers explained that they were more willing to send a consent request when the recording was meaningful to them, such as documenting a wedding, a family birthday, or a social gathering with close friends. In these situations, the personal value of keeping the video justified the effort of seeking approval.

> *[W1]"If I go to a cricket match [...] I would want that my audience's faces, who are watching me play cricket, who are watching me bat, their faces should have jubilation when I hit a six. Not just blurred faces. So, I would always prefer having [the original] video."*

(2) *Appropriateness of Recording.* Some wearers also mentioned that their willingness to send consent requests was contingent on whether they felt the recording was socially appropriate. If the filmed content was perceived as being invasive or inappropriate, they were reluctant to seek consent.

> *[W2]"If it's more of a humorous kind of video, then I would be actually a bit more cautious sending requests for consent. Like there are a lot of communication gaps in relationships and if somebody doesn't really like the way that they were like the target of some jokes or something, then [sending the request] will just be rubbing salt in their wounds [...] I wouldn't want that sort of thing."*

(3) *Consent Fatigue.* Wearers noted that the number of people in a recording strongly influenced their willingness to send requests. Asking for consent from a large group was perceived as impractical and burdensome by the wearers, leading them to avoid the process altogether. By contrast, if only a few people appeared in the frame, they were more inclined to reach out to those individuals, especially if they were well acquainted with them. Wearers not only perceived this burden for themselves but also reflected on it from the perspective of bystanders. They anticipated that receiving and responding to consent requests could feel inconvenient for others, particularly in social situations where people might not want to be interrupted. The social dynamics of consent requests also highly impact the usability of the protocol, we discuss these results in Section 6.3.

> *[W5]"You don't just take one picture, you take a lot of pictures, and then you select one, and maybe you want to take consent on one, and then you realize it's not that good. [...] so you want to replace it, and then you are again sending the request to everyone, and it'll be a hassle for them to accept it as well."*

6.1.2 How Bystanders Respond to Requests. Bystanders emphasized that their decision to grant consent was primarily influenced by the identity of the requester and the purpose of the recording.

Situational Context. An important consideration for bystanders was the situational context of the recording such as whether the setting was public or private, the content of conversations, and the duration of their presence. B9 further noted that their personal appearance or emotional state could also play an important role:

> *[B9] "Main factors I think would be who I am with, for example my friends or my family and whether I look good or not, whether I'm happy or not with myself, if I'm out with my family for example and we're just having like our own quality time and we don't wish to be recorded so yeah I think in that scenario [...], I would definitely decline the offer"*

Identity of Requester. The identity of the person requesting consent proved to be the most influential factor in participants' decisions. Trust and familiarity significantly lowered the threshold for granting consent, as B6 explained: *"If it's a friend then I will ask and if it's for memories then I think it's okay and would accept it. ... I would look at the context a little bit, but mostly I would accept it easily".*

Purpose of the Recording. The purpose of the recording was the most important factor participants considered when approving requests from strangers. B8 explained: *"I think they have to give reasons why they want to record me — if I don't know them. Obviously, I would think of why they want to do that. If I don't find a good reason, I won't give it to them".* B3 preferred wearers contact them through other channels before sending a consent request.

> *[B3]"I would feel that it would be best if they approached me through some other channel, like email or whatever, to let me know that I sent this request to them, so that I'm better prepared for accepting that request.*

6.1.3 Social Friction. A recurring theme in wearers' discussions of the consent mechanism was the potential for declined requests to create tension. While they acknowledged that recipients had every right to accept or deny a request, some foresaw that refusals could introduce awkwardness into their relationships, where a denial might be interpreted as a lack of trust or support.

> *[W4]"I think many a times even your partner will not allow you to use their face [on a video], and it'll be [...] a point of fight every other time: "Okay, what's wrong with you? Why would*

> *you not give me permission? I mean, do you not trust me". I think ultimately it'll come down to a trust thing."*

Bystanders reported mixed reactions when first encountering consent requests. B3 described their initial response as one of "surprise," "confusion," and "concern", whereas others interpreted the notification more positively:

> *[B5] "I think it will feel respectful for me that someone has recorded my video, even I didn't know (about recording), but he has taken a consent for the video. Then it will show some respect and some privacy concerns of that person."*

6.2 Perceived Efficacy and Limitations of Anonymization

Participants' assessments of obfuscation methods speak directly to bystander privacy needs (RQ1), revealing that effectiveness depends heavily on context and visual cues beyond facial features. Previous work has examined how obfuscation methods such as blurring and synthetic approaches are perceived in isolation [12, 27, 54, 55]. We present our findings both to confirm prior results and to report perceptions within the context of our protocol.

6.2.1 Blurring. Wearers and bystanders recognized the value of blurring as an obfuscation method, particularly in contexts involving strangers or incidental recordings. Participants also emphasized that blurring could be especially appropriate in contexts with cultural or religious sensitivities.

> *[W7]" Especially [for] women in Pakistan, they are like 'my face shouldn't be visible', so for that [blurring] is perfect."*

Other participants described it as the safest and most standard option for recordings in public, particularly when consent could not be obtained. Participants also mentioned that compared to existing indicators, such as the LED indicator on the Ray-Ban Stories, blurring was a more reliable and reassuring measure. They noted that it was particularly effective for keeping their identity anonymized when recordings were viewed by strangers. B5 emphasized its usefulness in public scenarios where *"someone has mistakenly or randomly taken a video"*.

6.2.2 Synthetic Face Replacement. Wearers and bystanders described synthetic face replacement as a more usable alternative to blurring, though their responses were nuanced. Both groups emphasized that its effectiveness depended on the realism of the generated face and the degree of similarity between the synthetic and original appearance.

Synthetic Attracts and Deflects Attention. Participants noted that the realism of synthetic faces could serve as a kind of deflection, making the modification less noticeable and deterring viewers from trying to guess the person's identity.

> *[B2]"It [synthetic] is preferred [compared to blurring] because when your face is blurred, people actually start looking at other things to identify who this person is. But when you have a face, you will just assume this is somebody [...]. So you won't really start making those connections that who this person actually is."*

At the same time, others felt that the same realism could draw unnecessary attention, with synthetic faces standing out in ways that distracted from the video.

> *[W2]"Actually applying a synthetic face on that might [put you] in frame of mind that "I know somebody who looks like this." We would have more information [in synthetic] than we would have had in the blurred video."*

Similarity Between Original and Synthetic Faces. Participants expressed strong views on the similarity of synthetic faces. W4 noted the feature as still being *"very recognizable"* believing that hyper-realistic face replacement would counterintuitively *"not preserve privacy at all"*, specifically in close-knit communities.

They expressed concern that synthetic faces could undermine anonymity unless the replacement was sufficiently distinct. Participants were concerned that the synthetic models might still retain features of the original face, which could compromise privacy. As B3 noted, *"AI replacement would probably try to fit something that's closer to my own face."*. Some bystanders expressed concern over how synthetic replacement could lead to the dissemination of misleading videos. B8 expressed concern that synthetic replacement could undermine the authenticity of the recording:

> *[B8] "If you do an AI thing, there are so many AI models and there's no one behind them. So, I think it (video) has no grounding."*

6.2.3 Personal and Non-Personal identifiers. Participants also identified two key limitations that could compromise anonymity.

Personal Identifiers. Participants noted that even with blurring or synthetic faces, people could still be recognized through personal identifiers such as hairstyle, mannerisms, voice, or body language.

> *[W4]" If I'm [at] a concert or in a party or somewhere in a social gathering [or] in a workplace, or [...] in a university where everybody knows everybody else [...] they will definitely figure out that this is you."*

> *[B7]"Often, the body structure of the person also gives away their identity. Which is actually quite common. You can identify the person without looking at the face. So, facial blurriness is not enough in that case and we can't even call this a corner case."*

Aside from visual obfuscation, participants identified the absence of audio privacy protections as a significant limitation of the system. They expressed concern that camera glasses could inadvertently capture sensitive conversations, potentially compromising privacy even when visual obfuscation was applied. Participants particularly noted that individuals familiar with the recorded person, such as friends or family members, could still identify them through personal vocal characteristics or speech patterns.

Non-Personal identifiers. Both wearers and bystanders also pointed out that non-personal identifiers like contextual elements and visual cues could serve as identifiers. B3 spoke of a "spectrum of paranoia," expressing anxiety that recognition might still occur "by some object that I'm holding, or the bag that I'm carrying, or something like that," even if their face was blurred.

> *[B3] “Some places might be recognizable as being very specific to an individual. So, for example, if there's a video recording of someone sitting in the dean's chair, then there's really no point blurring their face because in the dean's chair, you would probably only find the dean.”*

6.3 Usability of the Protocol

Wearer participants' reflections on protocol usability directly address RQ2, revealing three main factors that influenced their willingness to adopt opt-in privacy mechanisms. After experiencing the protocol through the interactive demonstration, wearers identified both benefits and challenges that shaped their perception of everyday use.

6.3.1 *Comfort Using the Protocol.*

Our findings show that the protocol gave some wearers a greater sense of comfort when recording in public.

Provides Freedom to Record. Wearers explained that the protocol made it easier to record continuously without worrying about inadvertently capturing random people in the background. They felt reassured that bystanders' privacy was being protected.

> *[W2]"With the protocol, the benefit would be that I could just keep it on all the time. I wouldn't have to be concerned about where I'm going [...] Things that we want to capture don't happen by plan—they happen by accident."*

Several wearers also highlighted that using the protocol made them feel good about recording, since they were actively preserving the privacy of those around them. W2 noted that the feature would *"make [them] feel more assured"* while recording. W4 mentioned that using the synthetic face replacement feature would make them *"satisfied in [their] heart"* knowing that they would not be breaching anyone's privacy. Other wearers commented on the capacity of the protocol bringing them at ease, as they felt an inherent guilt recording people while not knowing if they wanted to be recorded or not.

> *[W7]"I would probably feel more comfortable recording. Because I think even recording in public sometimes, like, I feel like I don't want to be recording someone if they don't want to be recorded. So this will be useful."*

Wearers highlighted that using the protocol provided them with a strong sense of moral and ethical reassurance. W7 explained that sending consent requests to family and friends would make them *"feel good about [themselves]"* framing it as a *"moral thing to do."*

Provides Social License. Wearers also framed the protocol as a means of providing them with social license to record in public. They explained that knowing obfuscation was in place allowed them to justify their actions to others. Some mentioned that if this protocol became common, it would increase the acceptability of these devices. Participants also mentioned that it could serve as a defense.

> *[W3]"I think it would give me enough of a defense. If I was wearing them in public, I would say don't worry, all videos and photos are automatically blurring faces."*

6.3.2 *Aesthetics.*

The aesthetics of the blurring and synthetic replacement mechanisms also impacted how wearers perceived the protocol's utility.

Blurring. Wearers described blurring as diminishing the overall aesthetics of recordings. They felt that blurred faces distracted from the content, drew unwanted attention, and were poorly suited for contexts such as social media where appearance and presentation mattered. Several noted that blurring removed the sense of connection to the video, creating faceless crowds that lacked personal touch or emotional resonance.

> *[W3]"For my personal recordings, I would prefer to have either the AI generated faces compared to just completely blurring [them] out because I think I'm losing a lot of information. I'm losing that personal touch and just the general environment that I'm in."*

Synthetic Face Replacement. Wearers frequently expressed a preference for synthetic face replacement over blurring, largely due to its enhanced visual appeal and suitability for public sharing. They described synthetic faces as more natural and engaging, making recordings feel realistic and better aligned with the aesthetics expected on social media.

> *[W7]"If [the video] is blurred, that video will look weird, the viewer's focus will also be distracted. But if it's a synthetic face, there will be eye movement, it will be interactive so when someone will watch it, they will watch it properly"*

6.4 Role of the Central Authority

The role of the central authority emerged as a consideration for both wearers and bystanders, with participants expressing divergent views on acceptable tradeoffs. While our protocol reduces reliance on a central authority by ensuring that recorded media is not shared, it still depends on that authority to deliver consent requests to bystanders. This introduces an important tradeoff where including more metadata in a consent request can help inform bystanders about the nature of the request, but it can also increase dependence on the central authority.

Our findings show participants have differing opinions on the exact nature of the metadata that should be included in the consent request with participants showing varying levels of trust. As a result, we leave the exact nature of the transferred metadata open for community debate. Some participants said they would trust the central authority completely if it was a big organization like Meta or Google. B8 mentioned how the central authority would provide authenticity to the consent protocol.

> *[B8]"I don't want to give this control to the person with the glasses. This should be monitored by the central authority because I'll be able to trust them"*

Other participants viewed the involvement of a central authority as a drawback and wanted to reduce reliance even further by limiting the amount of metadata shared while some viewed it as an acceptable compromise to maintain privacy 3. B7 compared the protocol to existing reliance on third party servers.

> *[B7]"If we're sending our images in WhatsApp (Messenger application), in end-to-end encrypted channels, and we trust that they don't know what we're sending. So, this is a small thing. So, compared to that, we have already taken major steps, which have far more significant consequences compared to this."*

6.5 Privacy Needs and Perspectives

Our findings show that wearers and bystanders have distinct privacy needs in the context of the protocol, directly informing both RQ1 (bystander privacy needs) and RQ2 (wearer usability requirements).

6.5.1 Wearer Perspectives. Many wearers reflected on the perspective of bystanders, noting that they too would want privacy protections in place if they were being recorded. This recognition shaped their expectations of the protocol, leading to suggestions for features and variations that would balance usability with privacy safeguards.

Choice and Context-Dependency. A significant number of wearers suggested that obfuscation should not be mandatory in every situation but rather an optional feature that they could activate depending on context. They valued the flexibility to decide when privacy protection was necessary. Several participants emphasized that this choice should be trusted to the user of the glasses, describing it as a moral and ethical right to know when to apply obfuscation and when it could be set aside.

Wearers proposed that the protocol should adapt to different recording contexts, such as private versus public spaces or professional versus casual environments. They believed that tailoring obfuscation and consent processes to the situation would make the system more practical and acceptable.

> *[W4]"It'll be great if there's a feature [...] if I'm in a private space, I can just do away with this whole thing [...] and just have a plain video [...] and not have to go through this process. [...] I can tell my glasses that this is a private event and it's okay, do not use that particular feature. And then when I'm outside, it comes down to your own personal, ethical, moral responsibility."*

This was contrasted by other wearer sentiments, for whom the convenience of recording ultimately triumphed over the need to preserve bystander privacy at all times. W8 made it explicit: *"If I have the option to apply that protocol on my own I will not do that"*. Other wearers suggested a pre-approval mechanism where trusted contacts, such as family or friends, could give permanent consent instead of being asked repeatedly. This was seen as a way to reduce social friction while still respecting the privacy of those not close to the wearer.

Synthetic Variation. Wearers expressed differing views on how synthetic faces should appear, weighing realism against clarity and transparency. Some preferred more natural and imperfect renderings to make recordings usable in social contexts. Others proposed slightly more animated or stylized faces, paired with visible indicators such as an AI-generated tag, to avoid confusion. Participants also cautioned against overly cartoonish designs, which they described as unpleasant and distracting.

6.5.2 Bystander Perspectives. Bystanders focused on the protocol's obfuscation methods and provided a range of different suggestions to improve the effectiveness of the blurring and synthetic replacement tiers. They advocated for more aggressive and comprehensive obfuscation strategies to be integrated into the protocol, such as a stronger blur filter, covering more area around the face, obscuring the entire person rather than just their face. B7 explicitly proposed the approach of blurring the entire person, *"It will be very necessary that you adopt this (person blurring) kind of aggressive anonymization." (B7)*

Bystanders also emphasized the need to address audio privacy. They proposed mechanisms such as masking and restoration or, in some cases, muting the audio altogether to prevent sensitive information from being exposed.

For synthetic face replacement, bystanders recommended making synthetic faces as different from the original as possible and even using AI to modify body language and mannerisms in order to reduce recognizability.

> *[B5]"If some gestures are changed by AI, then it will be good.[...] Body parts that can be recognizable by some group of people, not everyone, but some group of people. So if it is [...] feasible, we can change them as well, [...] and it will be more privacy."*

Bystanders also suggested improvements to the consent request mechanism. They mentioned that the wearer should provide a clear purpose for the request, especially when dealing with requests from strangers. Bystanders argued that this explanation should be compelling enough to justify the need for access. Some further proposed including an extended chat option to clarify questions and build trust. B1 mentioned how more context would impact their decision:

> *[B1]"If alongside the consent request, you could get the context like what do you need it for and if they give an option like a paragraph or a small sentence that I need this for this etc that would be better for the person accepting it... I think profiles should [...] ask them how much do you want to reveal, 'do I want to reveal my info while sending a request'. The option should be*

> *given to choose whether or not to share your name when sending a request. [...] I think if he revealed his name I would be more likely to accept it, otherwise not, otherwise less likely."*

6.5.3 Bystander Calls for Regulation. A striking and consistent finding was that all bystander participants supported the idea of making the protocol mandatory for manufacturers through regulation. As one participant explained, regulation was seen not simply as desirable, but as essential: *"This should definitely be there. If smart glasses are being worn without this protocol, I would want them banned." (B8)*

> *[B7] "Ideally, it (camera glasses) shouldn't have become mainstream. But again, it will happen. So, given that's the case, then why not? In fact, it should be a regulation."*

7 Discussion

Our findings reveal tensions between bystander privacy needs and wearer usability requirements in opt-in, privacy-by-default systems. Bystanders prioritize robust baseline obfuscation with contextually-appropriate consent mechanisms, while wearers value context-aware flexibility that adapts to social relationships and recording situations. We discuss how these findings address RQ1 and RQ2.

7.1 Wearer and Bystander Needs

Our findings reveal how wearers and bystanders conceptualize privacy in the context of opt-in PETs. This divide goes beyond simple preference differences and represents distinct frameworks for understanding privacy in the age of ubiquitous recording.

Privacy as a non-negotiable right. Bystanders approached the protocol from a rights-based perspective, viewing its protections as essential safeguards rather than optional features. Their unanimous call for mandatory implementation (Section 6.5.3) reflects a defensive stance towards an inherently invasive technology, aligning with previous work showing that bystanders experience discomfort and seek control when faced with ambiguous surveillance technologies [8]. Bystander demands for aggressive obfuscation, extending beyond faces (Section 6.5.2), furthers this stance.

Privacy as a contextual tool. Wearers evaluated the protocol through a social and practical lens. While they acknowledged the protocol's value in providing social license to record and reducing ethical burden (Section 6.3.1), they consistently desired to minimize its friction in everyday use. Wearer requests for features like pre-approved contact lists, context-aware obfuscation, and aesthetic improvements (Section 6.5.1) highlight that they view privacy mediation as something that should adapt to existing social dynamics rather than override them. This echoes prior research showing that wearers develop ad-hoc privacy negotiation strategies [16]. Our findings extend this work by demonstrating that wearers prefer context-aware technological interventions that complement these informal negotiations. Some wearers explicitly stated that they would not use privacy protections if given full control, revealing that convenience can trump ethical considerations for some participants.

These opposing frameworks suggest that successful privacy-mediating technologies must somehow reconcile rights-based and pragmatic perspectives without simply defaulting to restrictive approaches, which would limit adoption, or permissive ones, which fail to address legitimate privacy concerns.

7.2 Context as a Mediator

The role of context emerged as the most critical factor in how both groups evaluated the protocol, yet our findings reveal that context operates differently for wearers and bystanders. This divergence in understanding presents both a challenge and an opportunity for privacy-preserving design.

A lens through which to understand our findings is to compare how privacy negotiations occur within modern day contexts, such as smartphones, versus camera glasses. Nissenbaum's framework of contextual integrity posits that privacy depends on appropriate information flows within specific contexts [32]. Smartphone recording has evolved implicit interpersonal negotiations that, while not without friction, have become broadly accepted social norms: the visible act of pointing a phone, social cues indicating recording intent, and shared understanding of when recording is appropriate.

Camera glasses fundamentally disrupt these established norms, forcing bystanders into situations where they cannot determine if recording is occurring, let alone negotiate its appropriateness. This represents a violation of contextual integrity because the established mechanisms for negotiating information flows have been severed. While existing PETs acknowledge this breakdown, they largely propose solutions that themselves violate social norms, designs such as privacy bracelets, opt-out gestures, or broadcasting privacy preferences [22, 25, 34] all require bystanders to perform socially awkward actions that further disrupt, rather than restore contextual appropriateness. These rigid technical solutions fail because they attempt to impose new, unfamiliar negotiation mechanisms rather than building upon existing social practices.

Our protocol served as an enabler to explore how consent mechanisms might operationalize contextual integrity in practice. However, our findings reveal fundamental challenges. When wearers describe the hassle of sending consent requests to close friends and family (Section 6.1), they highlight how the protocol's mechanisms clashed with existing contextual norms where explicit permission-seeking would be socially inappropriate. In those intimate contexts, recording is often implicitly understood as acceptable, and formalizing consent introduces unwanted friction.

In contrast, bystanders' demands for justifications from strangers before granting consent (Section 6.1.2) reflect different contextual expectations. When the absence of a social relationship removes implicit trust, more explicit negotiation is required.

This tension suggests that future privacy-mediating technologies must move beyond uniform approaches and towards context-aware systems [53] that can distinguish relationships and situations. Relationship closeness, type of location (public vs. private), recording purpose, and group size all emerged as relevant factors in our results. However, a critical challenge remains: how can a system reliably identify these contextual factors and adjust privacy enforcement accordingly without itself becoming a privacy risk? Determining these factors requires additional data collection and inference that could introduce new privacy risks.

7.3 Design Directives for Opt-In Privacy

We discuss how our findings can be used to inform the design of opt-in, privacy-by-default mechanisms that take into consideration the differing perspectives of wearers and bystanders.

7.3.1 Context Dependent Application. Wearers seek contextual flexibility while bystanders require mandatory protection. Future systems should explore context-dependent ways of enabling or disabling the protocol i.e., the protocol could be relaxed in familiar private locations (e.g., a wearer's home) but mandatory in public spaces. Other factors could include social relationship (friends versus strangers) and real-time scene sensitivity inferred through on-device detection. This provides wearers with the flexibility to record in approved contexts while providing bystanders with mandatory protections in sensitive scenarios. The fundamental challenge is identifying the minimum contextual signals that allow meaningful adaptation without creating new privacy risks.

7.3.2 Mitigating Consent Fatigue. Meaningful consent inherently introduces fatigue for both wearers and bystanders. Future systems should allow bystanders to specify contextual constraints and preference settings. This allows bystanders to auto-accept or reject consent requests based on pre-defined settings like temporal scopes, location restrictions, purpose limitations, and friends/family lists. This reduces bystander fatigue while preserving agency and mitigates wearer concerns about inconveniencing bystanders.

7.3.3 Mitigating TTP-Associated Risks. Bystanders need context for informed decisions such as requester identity, recording purpose, and ability to ask questions; however, richer metadata expands TTP roles as privacy intermediaries. To balance these opposing considerations, progressive disclosure offers one path forward: consent requests start minimal and bystanders request additional context as needed, balancing information needs against data exposure. Future work can also explore decentralized architectures and examine whether these approaches can match centralized systems' utility while reducing privacy risks.

7.4 Directives for Camera Glass Manufacturers

We present a future direction for wearable privacy that enables more robust opt-in measures while still mediating the needs of both wearers and bystanders, and allowing for responsible use. Prior work has shown that the demand for stronger privacy mechanisms is not limited to bystanders [3, 4]. We present a framework that shifts the burden of privacy protection away from wearers by embedding privacy-preserving functions into the system, relieving them of the responsibility to individually negotiate social tensions. Current camera glass manufacturers [10, 18, 47] focus primarily on wearer usability while neglecting these social barriers that limit adoption. The presence of the protocol reassured both wearers and bystanders, making the act of recording more socially acceptable. Several wearers described how the protocol would allow them to justify recording in public spaces without appearing dismissive of privacy considerations. This points to an important value proposition for manufacturers: while the protocol introduces some operational overhead, it legitimizes recording practices and reduces the social stigma attached to camera glasses.

8 Conclusion

The growing popularity of camera glasses has raised significant privacy concerns. Existing PETs often prioritize privacy protection at the expense of recording usability, or conversely, maintain usability while offering weak privacy safeguards. At the same time, manufacturers largely focus on maximizing usability while overlooking the social tensions that shape the acceptability of camera glasses.

To address this gap, we conducted a qualitative user study examining the perceptions of wearers and bystanders regarding privacy-by-default opt-in mechanisms. We enabled our investigation by proposing a novel privacy-enhancing obfuscation protocol for camera glasses that provides mandatory on-device blurring, optional synthetic face replacement, and consent-based restoration.

Our findings revealed that bystanders express a strong need for mandatory opt-in obfuscation, whereas wearers acknowledged the importance of robust privacy but considered the consent-based mechanism inconvenient. Instead, wearers suggested context-based approaches in which the protocol could be adapted to the situation, as well as features such as pre-approved bystanders.

We also discussed how our findings highlight the complex dynamics of consent in privacy mediation. We propose that future PETs should attempt to balance the needs of wearers and bystanders by providing context-aware PETs that support effective consent negotiation.

References

[1] 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.* ACM, Singapore Singapore, 235–248. doi:10.1145/2906388.2906412

[2] Hafeez Ali A., Sanjeev U. Rao, Swaroop Ranganath, T. S. Ashwin, and Guddeti Ram Mohana Reddy. 2021. A Google Glass Based Real-Time Scene Analysis for the Visually Impaired. *IEEE Access* 9 (2021), 166351–166369. doi:10.1109/ACCESS.2021.3135024

[3] Divyanshu Bhardwaj, Alexander Ponticello, Shreya Tomar, Adrian Dabrowski, and Katharina Krombholz. 2024. In Focus, Out of Privacy: The Wearer's Perspective on the Privacy Dilemma of Camera Glasses. 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 577, 18 pages. doi:10.1145/3613904.3642242

[4] Taryn Bipat, Maarten Willem Bos, Rajan Vaish, and Andrés Monroy-Hernández. 2019. Analyzing the Use of Camera Glasses in the Wild. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3290605.3300651

[5] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3 (01 2006), 77–101. doi:10.1191/1478088706qp063oa

[6] Scott Carter and Jennifer Mankoff. 2005. When Participants Do the Capturing: The Role of Media in Diary Studies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, Portland Oregon USA, 899–908. doi:10.1145/1054972.1055098

[7] Soumyadeb Chowdhury, Md Sadek Ferdous, and Joemon M Jose. 2016. Lifelogging User Study: Bystander Privacy. In *Proceedings of the 30th International BCS Human Computer Interaction Conference.* BCS Learning & Development, Swindon, GBR, 3 pages. doi:10.14236/ewic/HCI2016.100

[8] 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.* ACM, Toronto Ontario Canada, 2377–2386. doi:10.1145/2556288.2557352

[9] Serge Egelman, Raghudeep Kannavara, and Richard Chow. 2015. Is This Thing On?: Crowdsourcing Privacy Indicators for Ubiquitous Sensing Platforms. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 1669–1678. doi:10.1145/2702123.2702251

[10] Google. 2025. Google Glass. https://www.google.com/glass/photography/. Accessed: 2025-09-07.

[11] Jonna Häkkilä, Farnaz Vahabpour, Ashley Colley, Jani Väyrynen, and Timo Koskela. 2015. Design Probes Study on User Perceptions of a Smart Glasses Concept. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. ACM, Linz Austria, 223–233. doi:10.1145/2836041.2836064

[12] Rakibul Hasan, Yifang Li, Eman Hassan, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. 2019. Can Privacy Be Satisfying?: On Improving Viewer Satisfaction for Privacy-Enhanced Photos Using Aesthetic Transforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–13. doi:10.1145/3290605.3300597

[13] Benjamin Henne, Christian Szongott, and Matthew Smith. 2013. SnapMe If You Can: Privacy Threats of Other Peoples' Geo-Tagged Media and What We Can Do about It. In *Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks*. ACM, Budapest Hungary, 95–106. doi:10.1145/2462096.2462113

[14] Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. 2017. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In *Advances in Neural Information Processing Systems (NeurIPS)*. Curran Associates Inc., Red Hook, NY, USA, 6629–6640.

[15] Bjørn Hofmann, Dušan Haustein, and Laurens Landeweerd. 2017. Smart-Glasses: Exposing and Elucidating the Ethical Issues. *Science and Engineering Ethics* 23, 3 (June 2017), 701–721. doi:10.1007/s11948-016-9792-z

[16] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. 2014. Privacy Behaviors of Lifeloggers Using Wearable Cameras. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, Seattle Washington, 571–582. doi:10.1145/2632048.2632079

[17] Hilary Hutchinson, Wendy Mackay, Bosse Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 17–24.

[18] isolomons. 2023. Introducing the New Ray-Ban | Meta Smart Glasses. https://about.fb.com/news/2023/09/new-ray-ban-meta-smart-glasses/

[19] Yahya Khawaja et al. 2026. Now You See Me, Now You Don't: Consent-Driven Privacy for Smart Glasses. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, Pisa, Italy.

[20] Yahya Khawaja, Shirin Rehman, et al. 2026. See me if you can (GitHub repository). https://github.com/SYSNET-LUMS/SmartGlassesPrivacy.

[21] Sam King and USENIX Association (Eds.). 2013. *22nd USENIX Security Symposium: August 14 - 16, 2013, Washington, D. C.* USENIX Association, Berkeley, Calif.

[22] Marion Koelle, Swamy Ananthanarayan, Simon Czupalla, Wilko Heuten, and Susanne Boll. 2018. Your Smart Glasses' Camera Bothers Me!: Exploring Opt-in and Opt-out Gestures for Privacy Mediation. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*. ACM, Oslo Norway, 473–481. doi:10.1145/3240167.3240174

[23] Marion Koelle, Matthias Kranz, and Andreas Möller. 2015. Don't Look at Me That Way! Understanding User Attitudes Towards Data Glasses Usage. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 362–372. doi:10.1145/2785830.2785842

[24] Mohammed Korayem, Robert Templeman, Dennis Chen, David Crandall, and Apu Kapadia. 2014. ScreenAvoider: Protecting Computer Screens from Ubiquitous Cameras. doi:10.48550/ARXIV.1412.0008

[25] Katharina Krombholz, Adrian Dabrowski, Matthew Smith, and Edgar Weippl. 2017. Exploring Design Directions for Wearable Privacy. In *Proceedings 2017 Workshop on Usable Security*. Internet Society, San Diego, CA, 6. doi:10.14722/usec.2017.23001

[26] 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. doi:10.1177/1541931213601694

[27] Si Liao, Hanwei He, Huangxun Chen, and Zhice Yang. 2025. Bystander Privacy in Video Sharing Era: Automated Consent Compliance through Platform Censorship. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–16. doi:10.1145/3706598.3713391

[28] Lin et al. 2015. Microsoft COCO: Common Objects in Context. arXiv:1405.0312 [cs] doi:10.48550/arXiv.1405.0312

[29] Meethu Malu and Leah Findlater. 2015. Personalized, Wearable Control of a Head-mounted Display for Users with Upper Body Motor Impairments. In *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 221–230. doi:10.1145/2702123.2702188

[30] Steve Mann. 1997. Wearable Computing: A First Step Toward Personal Imaging. *Computer* 30 (1997), 25–32. https://api.semanticscholar.org/CorpusID:28001657

[31] Meta. 2021. Introducing Ray-Ban Stories: First-Generation Smart Glasses. https://about.fb.com/news/2021/09/introducing-ray-ban-stories-smart-glasses/. Newsroom announcement.

[32] Helen Nissenbaum. 2009. *Privacy in context: Technology, policy, and the integrity of Social Life.* Stanford Law Books, USA.

[33] Yuqi Niu, Nicole Meng-Schneider, Weidong Qiu, and Nadin Kokciyan. 2025. "I Am Not the Primary Focus" - Understanding the Perspectives of Bystanders in Photos Shared Online. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–23. doi:10.1145/3706598.3713826

[34] Frank Pallas, Max-Robert Ulbricht, Lorena Jaume-Palasí, and Ulrike Höppner. 2014. Offlinetags: A Novel Privacy Approach to Online Photo Sharing. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems*. ACM, Toronto Ontario Canada, 2179–2184. doi:10.1145/2559206.2581195

[35] Yong Jin Park. 2015. Do men and women differ in privacy? Gendered privacy and (in)equality in the Internet. *Computers in Human Behavior* 50 (2015), 252–258. doi:10.1016/j.chb.2015.04.011

[36] Michael Quinn Patton. 2014. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice.* SAGE Publications, USA. Google-Books-ID: ovAkBQAAQBAJ.

[37] Robert A. Peterson and Dwight R. Merunka. 2014. Convenience Samples of College Students and Research Reproducibility. *Journal of Business Research* 67, 5 (May 2014), 1035–1041. doi:10.1016/j.jbusres.2013.08.010

[38] Raspberry Pi. 2019. Raspberry Pi 4 Model B Datasheet. https://datasheets.raspberrypi.com/rpi4/raspberry-pi-4-datasheet.pdf

[39] Rebecca S. Portnoff, Linda N. Lee, Serge Egelman, Pratyush Mishra, Derek Leung, and David Wagner. 2015. Somebody's Watching Me?: Assessing the Effectiveness of Webcam Indicator Lights. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 1649–1658. doi:10.1145/2702123.2702164

[40] Inc. Qualcomm Technologies. 2023. Snapdragon AR1 Gen 1 Platform Product Brief. https://docs.qualcomm.com/bundle/publicresource/87-86507-1_REV_B_Snapdragon_AR1_Gen_1_Platform_Product_Brief.pdf

[41] Nikhil Raina, Guruprasad Somasundaram, Kang Zheng, Sagar Miglani, Steve Saarinen, Jeff Meissner, Mark Schwesinger, Luis Pesqueira, Ishita Prasad, Edward Miller, Prince Gupta, Mingfei Yan, Richard Newcombe, Carl Ren, and Omkar M Parkhi. 2023. EgoBlur: Responsible Innovation in Aria. arXiv:2308.13093 [cs.CV] https://arxiv.org/abs/2308.13093

[42] Swati Rallapalli, Aishwarya Ganesan, Krishna Chintalapudi, Venkat N. Padmanabhan, and Lili Qiu. 2014. Enabling Physical Analytics in Retail Stores Using Smart Glasses. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*. ACM, Maui Hawaii USA, 115–126. doi:10.1145/2639108.2639126

[43] Philipp A. Rauschnabel, Alexander Brem, and Bjoern S. Ivens. 2015. Who Will Buy Smart Glasses? Empirical Results of Two Pre-Market-Entry Studies on the Role of Personality in Individual Awareness and Intended Adoption of Google Glass Wearables. *Computers in Human Behavior* 49 (Aug. 2015), 635–647. doi:10.1016/j.chb.2015.03.003

[44] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 2015. FaceNet: A Unified Embedding for Face Recognition and Clustering. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Boston, MA, USA, 815–823. doi:10.1109/CVPR.2015.7298682

[45] Shilpi Sharma and J. Sodhi. 2014. Implementation of Biometric Techniques in Social Networking Sites. *International Journal of Security and Its Applications* 8 (11 2014), 51–60. doi:10.14257/ijsia.2014.8.6.05

[46] Terence Sim and Li Zhang. 2015. Controllable Face Privacy. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Vol. 04. IEEE, Ljubljana, Slovenia, 1–8. doi:10.1109/FG.2015.7285018

[47] Seth Stevenson. 2016. Snapchat Releases First Hardware Product, Spectacles. http://www.wsj.com/articles/snapchat-releases-first-hardware-product-spectacles-1474682719

[48] Ivan E. Sutherland. 1968. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I (AFIPS '68 (Fall, part I))*. Association for Computing Machinery, New York, NY, USA, 757–764. doi:10.1145/1476589.1476686

[49] Eben Upton. 2013. *RPI 2003 Camera board.* https://www.raspberrypi.com/news/camera-board-available-for-sale/ Accessed: 2025-11-13.

[50] Rajesh Verma, Navdha Bhardwaj, Arnav Bhavsar, and Kewal Krishan. 2022. Towards Facial Recognition Using Likelihood Ratio Approach to Facial Landmark Indices from Images. *Forensic Science International: Reports* 5 (July 2022), 100254. doi:10.1016/j.fsir.2021.100254

[51] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing* 13, 4 (2004), 600–612.

[52] Yana Welinder. 2014. Facing Real-Time Identification in Mobile Apps & Wearable Computers.

[53] Maximiliane Windl, Petra Zsofia Laboda, and Sven Mayer. 2025. Designing Effective Consent Mechanisms for Spontaneous Interactions in Augmented Reality. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing System (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 18 pages. doi:10.1145/3706598.3713519

[54] Leslie Wöhler, Satoshi Ikehata, and Kiyoharu Aizawa. 2024. Investigating the Perception of Facial Anonymization Techniques in 360° Videos. *ACM Transactions on Applied Perception* 21, 4 (Oct. 2024), 1–17. doi:10.1145/3695254

[55] 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 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–16. doi:10.1145/3613904.3642103

[56] Zhiliang Xu, Zhibin Hong, Changxing Ding, Zhen Zhu, Junyu Han, Jingtuo Liu, and Errui Ding. 2022. MobileFaceSwap: A Lightweight Framework for Video Face Swapping.

[57] Roberto Yus, Primal Pappachan, Prajit Kumar Das, Eduardo Mena, Anupam Joshi, and Tim Finin. 2014. Demo: FaceBlock: Privacy-Aware Pictures for Google Glass. In *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '14)*. Association for Computing Machinery, New York, NY, USA, 366. doi:10.1145/2594368.2601473

[58] Lan Zhang, Xiang-Yang Li, Wenchao Huang, Kebin Liu, Shuwei Zong, Xuesi Jian, Puchun Feng, Taeho Jung, and Yunhao Liu. 2014. It Starts with iGaze: Visual Attention Driven Networking with Smart Glasses. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*. ACM, Maui Hawaii USA, 91–102. doi:10.1145/2639108.2639119

[59] Ziyang Zhang, Chong Bao, Xiaokun Pan, Chia-Ming Chang, Takeo Igarashi, and Guofeng Zhang. 2025. Through the Lens of Privacy: Exploring Privacy Protection in Vision-Language Model Interactions on Smart Glasses. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3706599.3720234

[60] Yuhang Zhao, Yaxing Yao, Jiaru Fu, and Nihan Zhou. 2022. "If Sighted People Know, I Should Be Able to Know:" Privacy Perceptions of Bystanders with Visual Impairments around Camera-based Technology. arXiv:2210.12232 [cs] doi:10.48550/arXiv.2210.12232

[61] Niek Zuidhof, Somaya Ben Allouch, Oscar Peters, and Peter-Paul Verbeek. 2024. Perspectives on the Acceptance and Social Implications of Smart Glasses: A Qualitative Focus Group Study in Healthcare. *International Journal of Human–Computer Interaction* 40, 2 (Jan. 2024), 149–159. doi:10.1080/10447318.2022.2111046

Appendix

Table 4: Semi-structured interview guideline for wearer and bystander sessions.

Part	Wearer Protocol	Bystander Protocol
1. Introduction & Oral Consent	**Introduction & Oral Consent** Thank you for your participation in our study on smart glasses. Today, we will talk about your experiences and show you different ways that recordings can be processed. Everything will be kept confidential. You may stop at any time. Do you consent to continuing? Do you consent to recording this interview? *Notes:* Begin audio recording after verbal consent.	*Same questions as wearer.*
2. Warm-Up Questions	Could you tell me about your general experience with the smart glasses over the past X days? *Notes:* Open-ended. Use probing questions if needed.	Have you ever encountered someone using smart glasses before this study? How did you feel when you first saw someone wearing them during this study? Did anything surprise you about the experience of being around smart glasses? *Notes:* Use as a gentle entry point before showing obfuscation features.
4. Features	**Tier 1 – Automatic Blurring** What comes to mind when you see the blurred video? How do you view the privacy aspects of this approach? In what situations do you think blurring might be most or least applicable? What would your experience be like using this feature? How might this feature influence your recording behavior? **Tier 2 – Synthetic Face Replacement** What comes to mind when you see the synthetically replaced video? How do the synthetic faces relate to the original people in the recording? How do you view the privacy aspects of this approach? In your view, how effective is this approach compared to blurring? In what scenarios would this method be most relevant for you? How might this feature influence your recording behavior? **Tier 3 – Consent Requests** What are your thoughts on this consent feature? How do you see this addressing privacy considerations? What would influence your likelihood of using it in different situations? What considerations would shape your decision to send or decline to send a request? How would you respond to having a request declined?	**Tier 1 – Automatic Blurring** What comes to mind when you see the blurred video? How do you view the privacy aspects of this approach? In what situations do you think blurring might be most or least applicable? What would your experience be like using this feature? How might this feature influence your interactions in situations where people wear smart glasses? **Tier 2 – Synthetic Face Replacement** What comes to mind when you see the synthetically replaced video? How do you view the way it represented your appearance? How do you view the privacy aspects of this approach? In your view, how effective is this approach compared to blurring? In what scenarios would you choose replacement, blurring, or neither? **Tier 3 – Consent Requests** What are your thoughts on this consent feature? How would you characterize receiving a consent request? How do you see this addressing privacy considerations? What would influence your likelihood of using it in different situations? What considerations would influence your decision to approve or deny? What would your experience be with accepting or rejecting requests?
5. Overall Protocol / Privacy & Central Authority	What are your thoughts on the overall approach? How do you view the balance between wearer needs and bystander privacy in this system? This system would be managed by a central authority. What are your thoughts on having such an authority manage this system? What is your perspective on how they might handle your data? What changes or additions would you consider for this system?	*Same questions as wearer.*
6. Improvements / Final Thoughts	What issues or concerns do you see with this system? What improvements would you suggest? Do you have any final thoughts or feedback?	*Same questions as wearer.*
7. Outro & Debriefing	Thank you for your participation. Do you have any final questions? Please remember not to share details of the protocol with others yet. You will now receive your compensation. *Notes:* Stop recording. Provide compensation.	*Same questions as wearer.*