



Figure 1: The MirrorCam prototype is a chest-worn camera including a screen-based status indicator. We contribute our experiences from an in-the-wild study where participants tested the prototype in their everyday lives and collected 79 hypersubjective impressions in their pen-and-paper diaries.

KEYWORDS

Wearable computing; body-worn cameras; social acceptability; field study; diary study.

Evaluating a Wearable Camera's Social Acceptability In-the-Wild

Marion Koelle

University of Oldenburg
Oldenburg, Germany
marion.koelle@uol.de

Wilko Heuten

OFFIS – Institute for IT
Oldenburg, Germany
wilko.heuten@offis.de

Torben Wallbaum

OFFIS – Institute for IT
Oldenburg, Germany
torben.wallbaum@offis.de

Susanne Boll

University of Oldenburg
Oldenburg, Germany
susanne.boll@uol.de

ABSTRACT

With increasing ubiquity, wearable technologies are becoming part of everyday life where they may cause controversy, discomfort and social tension. Particularly, body-worn “always-on” cameras raise social acceptability concerns as their form factors hinder bystanders to infer whether they are “in the frame”. Screen-based status indicators have been suggested as remedy, but not evaluated in-the-wild. Simultaneously, best practices for evaluating social acceptability in field studies are rare. This work contributes to closing both gaps. First, we contribute results of an in-the-wild evaluation of a screen-based status indicator testing the suitability of the “displayed camera image” design strategy. Second, we discuss methodical implications for evaluating social acceptability in the field, and cover lessons learned from collecting hypersubjective self-reports. We provide a self-critical, in-depth discussion of our field experiment, including study-related behavior patterns, and prototype fidelity. Our work may serve as a reference for field studies evaluating social acceptability.

CHI'19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland Uk

© 2019 Copyright held by the owner/author(s).

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI'19 Extended Abstracts), May 4–9, 2019, Glasgow, Scotland Uk*, <https://doi.org/10.1145/3290607.3312837>.

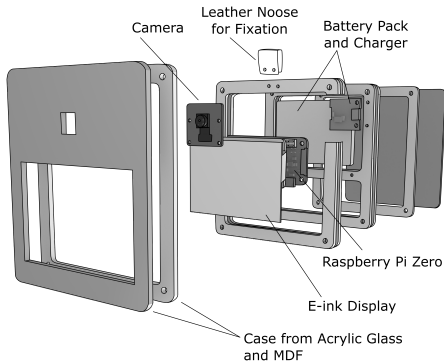


Figure 2: During the two-day field test, our participants wore the MirrorCam prototype in various locations and situations, such as public transport (top), and university (bottom). They collected bystander feedback in their diary.

INTRODUCTION & PRIOR WORK

The social acceptability of an interface commonly includes two perspectives: the user, and the observer (or bystander) [11]. Body-worn cameras cause that the observer – being in-view of the camera – also becomes the observed. In addition, contemporary body-worn cameras often do not sufficiently indicate whether they are “ON” or “OFF”, and who is within their field of view. A lack of notice may result in a lack of situation awareness on the bystanders’ side, and a lack of justification on the user’s side [7]. A potential remedy is to announce information about the device and its field of view to bystanders by displaying the camera’s image; a strategy which is utilized by body cams used for policing, but has not yet been evaluated in a broader context. In this work, we investigate the potential of screen-based status indicators for casual usage based on a collection of 79 diary entries. Each of the nine study participants wore the MirrorCam prototype (depicted in Figure 1), for two subsequent days in their everyday life, collecting self-reported, hypersubjective impressions and bystander feedback. We discuss and analyze both, experiment and outcome, and provide practical, methodical implications for evaluating social acceptability in the field.

Candid and revealed interactions. While mobile and wearable computing mostly aimed to design interactions with devices and interfaces as unobtrusive or inconspicuous as possible, some approaches advocate more “candid” interactions [3]. Such candid, i.e., revealed or amplified, interactions leverage situation awareness on the observer’s side by explicitly pointing out core motives (e.g., application type or purpose) of the interaction with a device. In an early work, Bellotti et al. [1] employed this principle to provide bystanders with information about a stationary camera; from a display mounted next to the camera they could obtain feedback about the captured imagery, whether they are in range and how they look like. This “Confidence Monitor” is described as trustworthy, meaningful and appropriately timed, but – being stationary – was not transferred to wearable computing devices. Utilizing (additional) displays to achieve this kind of transparency (or “candidness”) for wearables has mainly been explored in the context of virtual reality. To reveal social signals and leverage communication between the headset’s user and bystanders, researchers proposed to augment Virtual Reality headsets with one or more screens facing the bystander. These might overlay the user’s occluded eye-movements [2], or the virtual environment (s)he is in [2, 4], or let bystanders “see-through” the headset by displaying 3-dimensional renderings of the user’s face [4]. In previous work, Koelle et al. [7] explored design strategies for status indicators for body-worn cameras based on 8 low-fidelity artifacts originated from co-design sessions. Four of these artifacts proposed to display the camera image (or a derived abstraction) as status indicator, which was rated by experts as well understandable and intuitive, but not tested in-the-wild. With our present work we close this gap by testing out the potential of screen-based status indicators for wearable cameras in-the-wild.



Processing Unit: Raspberry Pi Zero v1.3 including camera module v2.1 with a resolution of 1080p@30 fps. A 2000mAh battery pack, and a 3.7V to 5V converter module provide it with power supply for approx. 6h.

Display: PaPiRus driver board with a 2.7" eInk/ePaper screen; screen resolution: 264 x 176 pixels, screen size: 60mm x 40mm. In contrast to a colored LCD or TFT display, this allows for a wider viewing angle, lower power consumption, and is less prone to sunlight reflection.

Case: the 90mm x 105mm x 33mm case is composed of 3mm acrylic glass, and 3mm medium-density fiberboard (MDF) using a sandwich technique. Layers are fastened with M2.5 screws, a leather noose is screwed to the top delimiter for safe fixation to a lanyard.

Sidebar 1: Technical details of the MirrorCam prototype

Contributions. We contribute a discussion of methodical implications for evaluating social acceptability in the field, based on a diary study testing the social acceptability of the “displayed camera image” design strategy. In particular, we cover lessons learned from collecting hypersubjective self-reports using diary studies, problems arising from prototype fidelity, and issues with study-induced user behavior. What we intentionally, and unintentionally learned from the presented experiment may serve as reference for other researchers conducting studies on social acceptability in the field.

METHOD

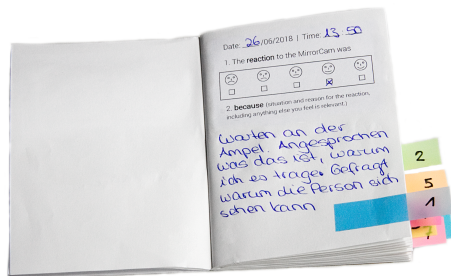
Diary studies are beneficial to gather insights about technology phenomena in uncontrolled, in-the-wild settings, and limit the impact of retrospective interpretation [9]. Using a dedicated prototype (here: MirrorCam, technical details in Sidebar 1.) as research vehicle, we conducted two day field trials (“diary study”, Figure 2), where the participants noted down their experiences in a diary.

Enrollment and study procedure. We recruited our participants on campus, as students – in contrast to professionals – are more mobile during a typical work day. This allowed to maximize the variety of locations and situations explored during the study, while minimizing the time effort for each participant. They registered for a period of four consecutive days, with intervals being spread out to cover all weekdays, including weekends. The enrollment procedure followed Hoyle et al. [5], with a 30 min briefing session (Day 1), a two days field trial (Days 2+3), and a 30 min debriefing session (Day 4). During the field trial the participants were asked to wear the MirrorCam prototype whenever possible. They were asked to wear the prototype in a way that its screen was observable by those around, explain the purpose of study and device to persons in their vicinity, and collect their reactions in the diary. They also received a set of information cards to be handed out to third-parties inquiring about the study (including a link to FAQs). In contrast to [6] our prototype did not persistently store (image) data, and did not require in-situ delete. Each participant received a 20€ Amazon voucher.

Data collection and analysis. During the field trial (Days 2+3), the participants wrote down their impressions as well as bystander’s reactions into their diary. Each diary page contained a 5-pt Kunin Scale [8], and space for a free-text explanatory statement, as depicted in Figure 3. After the field trial, the participants were re-invited to the lab (Day 4) for a semi-structured interview, where they discussed their overall experience. In addition, they were asked to point out diary entries that they found noteworthy, surprising, or most relevant. The interview was audio-recorded. Both, diary entries and interviews, were transcribed, digitized and qualitatively analyzed using inductive category development [10]. In their diaries, our participants not only report individual views, but also feedback and reactions they witnessed during the field trial: the collected qualitative data is not only subjective, but hypersubjective. Thus, we count themes (denoted as *n*) based on the number of diary entries (*N*=79), and denote participants as *P*, diary entries in verbatim, and interview excerpts in *italic*.



Participants collected their impressions in a A7 diary. During the debriefing interview selected diary entries were discussed and marked for reference (page markers).



Each page consisted of fields for date and time, a 5-pt Kunin Scale: “The reaction to the MirrorCam was...”, and space for a free-text explanation, including the (given) reason for the reaction and situation: “because... (situation and reason for the reaction, including anything else you feel is relevant.)”.

Figure 3: Pen-and-paper journal used by the participants to record their impressions during Days 2+3 of the study.

¹The list of Do’s and Dont’s, based on [6], is included with the supplementary material.

FINDINGS

Nine participants (4f, 5m), aged 22 - 30 ($M=26$, $SD=2$) collected $N=79$ distinct diary entries, recording between 2 and 18 ($M=9$, $SD=6$) entries each. They showed surprise to having received mostly neutral ($Mdn=3$, $SD=1$) feedback. One fourth of the diary entries report expectations of a (positive or negative) reaction, where the participants did not perceive any (no reaction, $n=20$). P4 puts down: “Quick glances (at most)”, which resonates with the participants’ self-reports during debriefing: “even when I was roaming university campus” (P7).

Wearing a camera in public. Many bystanders displayed curiosity and interest ($n=17$). However, multiple entries also report avoidance behavior ($n=6$) and skepticism ($n=9$). In particular, participants reported that “recording” in terms of persistent data storage was key: “[I told him] it doesn’t record,[...] then he was like ok, then it is somehow interesting; tell me more. Why do you do this [...] there is no use if it does not store anything” (P4). These show the relevance of the persistence of recorded imagery to social constraints (recording matters, $n=8$). On the other hand, diary entries also showed evidence of a perceived culture of surveillance: “I’m under surveillance anyway” (P3).

(Mis-)interpreting the screen. The integrated screen was noted positively for increasing transparency ($n=3$), and sparking conversations. Explicitly positive reactions ($n=5$), such as bystanders referring to the device as “cool” or “funny”, and waving at it were also noted: “[They] stood in front of it and waved [at it]. So like ‘Hey what’s that? Hey I can see myself’” (P6). Only twelve entries report bystanders recognizing themselves in the integrated screen (recognition, $n=12$), but rather that the prototype was misinterpreted, e.g., as game or jewelry, and not being recognized as camera (interpretation, $n=7$). Two participants (P3, P5) observed that bystanders understood the prototype as assistive technology and attributed this to its single piece appearance, and medium fidelity (c.f., Excerpt 1).

Self-perception of the participants. Though mostly targeting bystander reactions, many entries reflect the participants’ self-perception. The felt observed or looked at, even without any explicit reaction to the prototype (perceived attention, $n=13$). P7 questions their objectivity “So I guess there was a discrepancy between my perception and how it actually was, because I think people didn’t actually look at me [...]”. In addition, participants stated to have enrolled to test their self-confidence; Some expressing surprise about their reluctance during the study (c.f., Excerpts 2 and 3).

Study-related behavior patterns As illustrated by Excerpt 4, three participants explicitly sought bystander feedback to achieve a high number of diary entries. They reported to have chosen clothing, locations and/or body postures that highlighted them wearing the device. In contrast, as also reflected by the number of entries’ high inter-subject variance, others wore the prototype only where they felt confident to not trigger concerns. In addition to the study’s Do’s and Dont’s¹, some participants took additional measures to not be accused of surreptitious picture-taking, e.g., by acquiring consent from a supermarket’s branch manager before entering the store (P6).

Excerpt 1: Interpreting the prototype as assistive technology.

He also said, that his first impression was somehow as if I had some handicap, and that it [the prototype] was a support for it. Maybe in a way that it records what I see here and then somehow gives me an input. I think that's why people tended to look so nervous and embarrassed at me.

-P3

Excerpt 2: Study participation as a test of self-confidence.

I have to say that I found it surprisingly difficult to wear the thing, not because it was uncomfortable or so, but because it has cost me quite some effort to take with me in social situations. I had expected this a little, that's why I found the study so exciting, because I thought, ok, I wanted to test myself, too. [...] I also react to other people wearing cameras [...] a situation that I do not find so pleasant. So it was definitely fascinating.

-P4

Excerpt 3: Feeling comfortable by avoiding conflicts.

I had difficulties on the first day and also a little bit on the second day to find any situations where I felt comfortable using [the wearable camera].

-P8

DISCUSSION

In the following, we discuss both results, and methodical implications of our diary study.

Value and limitations of diary studies Diary studies (contrary to observations, or lab experiments) induce typical limitations (c.f., Lazar et al. [9]): participants may not follow through, and only sparsely record entries, plus recruitment is more slow-going as with a less intrusive study (e.g., a survey). In addition, self-selection bias can occur, as volunteers likely have greater technology affinity than average. However, this is not unrealistic: tech-savvy audiences are also more likely to become early adopters, and thus, encounter similar reactions in public as our participants did. Moreover, our study design anticipates, and partially mitigates effects of self-selection bias due to the hypersubjective nature of the reports: participants also recorded how they were perceived by others. Such (hyper)subjective reports from diary studies as ours can provide valuable insights on social effects of technology, but, being subjective, have to be taken with care.

Mitigating study-related behavior While avoidance of (negative) reactions is likely to reflect real usage patterns, some participants might feel the need to provoke as many reactions, i.e., diary entries as possible. These might be biased or unrealistic. A remedy might be to recording the nature of the participants' behavior and encounters: were they acting outgoing, reserved, or provocative? Was this behavior characteristic or atypical for themselves? What was their relationship to the bystander(s)? Hence, in addition to the measures taken in our study, careful one-on-one briefing, oral explanation of the diary entries, and equal pay for all participants, future work should further contextualize the participant's self-reports to account for potential effects from study-related behavior.

Recording vs. not recording. For ethicality, our prototype did not persistently store image data, which raised questions about the "value" for the participants: why would they wear such a device, if they do not get to keep the images? This might be problematic, as perceived utility (c.f., Profita et al. [12]) can influence social acceptability. Thus, it might be sensible to introduce "added value" (e.g., images to keep, an app or game) for the participants to increase realism in future studies. On the other hand, our results also indicate that persistently storing, in contrast to "piping-through" imagery, does affect potential privacy concerns, and thus, transitively, social acceptability, which has implications for technologies using a camera as sensors, e.g., image-based tracking. Future work on status indicators would thus not only aim to communicate what data is captured, but also what for.

Prototype fidelity. While the MirrorCam prototype was perceived less salient than its size might have suggested, its medium fidelity also had unforeseen effects, as it created the impression of the prototype being an assistive device. As assistive devices tend to be more accepted than consumer "just-for-fun" devices [12], such "AT-Effects" might bias social acceptability studies with non-consumer devices. Thus, Future work should consider to what extent social factors can be evaluated with low-fi prototypes outside of lab environments (where participants "imagine" the final interface).

Excerpt 4: Provoking bystander feedback.

At the beginning I felt a bit weird, so when I went out for the first time I thought, hm, usually no one really looks at me [...] so I just wore a black T-shirt, where I thought I'd feel a bit comfortable with [the black camera]. Then, I went for a short round outside, ran a few errands, and then I put on a white T-shirt (laughs) and thought, so now I want people to recognize it, because there was not really any content, people have looked somehow, but no one had reacted [to the camera]. [...] At first, I have also been wearing my hair down, then I made a pony tail, in order to present it [the camera] more, because I wanted to have some feedback [...] so a bit more provocative.

-P6

CONCLUSION & FUTURE WORK

We presented results and experiences from a field test of a wearable camera with a screen-based status indicator, which was noted positively for increasing transparency, but not always recognized by bystanders. Our findings furthermore indicate that (1) diary studies are suitable means for evaluating aspects of social acceptability, and collecting hypersubjective impressions, but that (2) studies investigating social acceptability aspects should account for “perceived utility”. Thus, equip the to-be-tested device with an “added value” for the participants. (3) Prototype fidelity may impact on bystanders’ reactions and interpretations. Future work might provide methods and best practices to mitigate such effects in social acceptability studies, e.g., by employing Wizard-of-Oz techniques.

REFERENCES

- [1] Victoria Bellotti and Abigail Sellen. 1993. *Design for Privacy in Ubiquitous Computing Environments*. Springer Netherlands, Dordrecht, 77–92. https://doi.org/10.1007/978-94-011-2094-4_6
- [2] Liwei Chan and Kouta Minamizawa. 2017. FrontFace: Facilitating Communication Between HMD Users and Outsiders Using Front-facing-screen HMDs. In *MobileHCI '17*. ACM, Article 22, 5 pages. <https://doi.org/10.1145/3098279.3098548>
- [3] Barrett Ens, Tovi Grossman, Fraser Anderson, Justin Matejka, and George Fitzmaurice. 2015. Candid Interaction: Revealing Hidden Mobile and Wearable Computing Activities. In *UIST '15*. ACM, 467–476. <https://doi.org/10.1145/2807442.2807449>
- [4] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *CHI '18*. ACM, Article 54, 13 pages. <https://doi.org/10.1145/3173574.3173628>
- [5] Roberto Hoyle, Qatrunnada Ismail, David Crandall, and Apu Kapadia. 2015. Challenges in Running Wearable Camera-Related User Studies. In *CSCW Workshop: The Future of Networked Privacy: Challenges & Opportunities*. Vancouver, BC, Canada. https://networkedprivacy2015.files.wordpress.com/2015/02/hoyle_etal-cscw2015.pdf
- [6] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. 2014. Privacy Behaviors of Lifeloggers Using Wearable Cameras. In *UbiComp '14*. ACM, 571–582. <https://doi.org/10.1145/2632048.2632079>
- [7] Marion Koelle, Katrin Wolf, and Susanne Boll. 2018. Beyond LED Status Lights - Design Requirements of Privacy Notices for Body-worn Cameras. In *TEI '18*. ACM, 177–187. <https://doi.org/10.1145/3173225.3173234>
- [8] Theodore Kunin. 1955. The Construction of a New Type of Attitude Measure. *PERS PSYCHOL* 8, 1 (1955), 65–77.
- [9] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research Methods in Human-Computer Interaction*. Morgan Kaufmann.
- [10] Philipp Mayring. 2014. *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. AUT.
- [11] Calkin S. Montero, Jason Alexander, Mark T. Marshall, and Sriram Subramanian. 2010. Would You Do That?: Understanding Social Acceptance of Gestural Interfaces. In *MobileHCI '10*. ACM, 275–278. <https://doi.org/10.1145/1851600.1851647>
- [12] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K. Kane. 2016. The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use. In *CHI '16*. ACM, 4884–4895. <https://doi.org/10.1145/2858036.2858130>