Ubiquitous Machinery Monitoring – A Field Study on Manufacturing Workers' User Experience of Mobile and Wearable Monitoring Apps

SEBASTIAN MÜLLER, Eastern Switzerland University of Applied Sciences, Switzerland
MATTHIAS BALDAUF, Eastern Switzerland University of Applied Sciences, Switzerland
ARNE SEELIGER, ETH Zurich, Switzerland and Geberit Produktions AG, Switzerland

Fig. 1. In a four-week field study, we investigated manufacturing workers' user experience of mobile and wearable apps for ubiquitous machinery monitoring in modern production facilities under productive operation.

Fueled by ongoing digitization efforts, manufacturing is currently undergoing a transformational process towards interconnected machinery and workforce, which enables a wide range of interactive monitoring and controlling applications. Whereas existing user-centered work addressed remote monitoring from office workplaces, it remains unclear how manufacturing workers experience and adopt machinery monitoring apps on mobile and wearable devices. To close this gap, we conducted a four-week field study in a running factory to study workers' overall user experience and acceptance of such monitoring apps, the subjective impact on their work routines, and their preferred device type. Under productive operation, 11 manufacturing workers used functional application prototypes on smartphones and smartwatches to receive notifications of machine incidents. In 22 individual interviews and two focus groups, we collected the participants' impressions and assessments. Based on these results, we derive a set of recommendations for designing and deploying machinery monitoring apps for manufacturing workers.

CCS Concepts: • **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**; **Field studies**.

Additional Key Words and Phrases: Smart factory, wearable, worker, monitoring, automation

198

Authors' addresses: Sebastian Müller, sebastian.mueller@ost.ch, Eastern Switzerland University of Applied Sciences, Rosenbergstrasse 59, St.Gallen, Switzerland, 9001; Matthias Baldauf, matthias.baldauf@ost.ch, Eastern Switzerland University of Applied Sciences, Rosenbergstrasse 59, St.Gallen, Switzerland, 9001; Arne Seeliger, aseeliger@ethz.ch, ETH Zurich, Weinbergstrasse 56/58, Zurich, Switzerland, 8006 and Geberit Produktions AG, Schachenstrasse 77, Jona, Switzerland, 8645.

2573-0142/2022/9-ART198
https://doi.org/10.1145/3546733

ACM Reference Format:
Sebastian Müller, Matthias Baldauf, and Arne Seeliger. 2022. Ubiquitous Machinery Monitoring – A Field Study on Manufacturing Workers' User Experience of Mobile and Wearable Monitoring Apps. *Proc. ACM Hum.-Comput. Interact.* 6, MHCI, Article 198 (September 2022), 22 pages. https://doi.org/10.1145/3546733

1 INTRODUCTION

Many industrial companies are currently undergoing a transformational process from conventional production procedures to flexible and smart manufacturing [25]. This change is enabled through networked physical machinery and the adoption of a variety of novel technologies [24]. Early research addressed the topic mainly from a technological point of view, thereby giving limited consideration to the human factor. More recently, the role of the human within smart factories has been emphasized as an important topic for both academia and industry. Specifically, workers and machine operators are confronted with an increasingly complex and dynamic work environment [16]. In this context, the intervention by humans to troubleshoot or prevent machine failure has been identified as an important type of task for operators and shop-floor workers [15].

Existing research that investigates the support of workers and machine operator falls into two categories. On the one side, visualization techniques have been specifically tailored to industrial applications such as production [39]. These approaches mainly rely on large screens to present the plethora of real-time process data acquired from modern production lines. Machine operators and workers, however, perform most of their work on the shop-floor. Hence they only have limited access to large, stationary screens.

On the other side, wearable and mobile assistance solutions for industrial use cases have received increased attention. This development has been driven partly by the interest in wearable, usually head-worn, augmented reality (AR) systems (see e.g., [26, 27, 30, 31]). Besides this specialized hardware, the usage of more common consumer market technologies, like smartphones or smartwatches, has been identified as potentially suitable for industrial applications [15]. In this context, existing research provides first insights into information presentation on such devices [17, 37]. Further research in this area, however, is scarce and existing work often lacks sufficient grounding in industrial applicability. Specifically, prior works addressed only selected sub-tasks of shop-floor workers (e.g., quality control [37]), which disregards the variety and complexity of their work. Existing studies also tend to rely on lab-based settings, often incorporating untrained participants (e.g., [17]). Finally, aspects of user experience are often not addressed in sufficient detail due to the usage of high level questionnaires (e.g., [2]). A comprehensive assessment of how manufacturing staff experiences the usage of mobile and wearable assistance systems in their daily work is therefore missing.

We bridge this gap by evaluating the user experience of ubiquitous machinery monitoring in form of mobile and wearable apps (Figure 1) in state-of-the-art production facilities. Over a period of more than four weeks, we equipped eleven shop-floor workers with mobile and wearable devices for machine monitoring and conducted both individual interviews and focus groups. We therefore contribute to the existing body of research in two ways.

(1) We present findings on the user experience of mobile and wearable machine monitoring apps from an extensive field study with professionals.
(2) Based on the results of the field study, we provide a set of recommendations for designing and deploying machinery monitoring apps for factory workers.

2 RELATED WORK

Our work builds upon previous research in the areas of (1) smart devices in industrial settings, especially wearable and mobile ones, (2) machine monitoring in the smart factory, and (3) usability aspects of industrial wearable and mobile devices.

2.1 Smart Devices in Industrial Settings

Many industrial firms are striving for more flexible and efficient processes by leveraging recent advancements in the field of information technology, data analysis, or sensor technology [25]. As part of this transformation process, smart and networked devices have received particular attention from both industry and academia. Smart industrial wearable and mobile devices are usually empowered with computing and communications capabilities in order to support connectivity and intelligent applications [29]. Generally, such devices feature both human interaction and data collection functionalities [14]. In spite of these similarities, smart industrial wearable and mobile devices span a great range of formats. For instance, previous work discussed smart watches, smart phones, smart glasses, or smart gloves. Mark et al. [18] provide a detailed overview of various devices that can be regarded as smart industrial devices for worker assistance.

A large body of research focuses on AR devices. For instance, Zheng et al. [38] present a wearable system based on Google Glass. Url et al. [31] as well as Seeliger et al. [27] employ Microsoft's HoloLens in a production setting. Likewise, Jetter et al. [12] utilize a mobile AR tool for automotive maintenance. A more detailed review of AR solutions in manufacturing can be found in the work of Nee et al. [20]. Beyond AR applications, only few works consider smart mobile devices. For instance, Zenker and Hobert [37] investigate the usage of smartwatches in an industrial setting. Along similar lines, Paelke et al. [22] discuss user interfaces for cyber-physical systems, thereby emphasizing the hands-free interaction with smartwatches. The authors, however, do not provide any comparative evaluation in the field. Overall, there is limited research on smart mobile devices in industrial settings and existing research calls for critical studies on the subject matter [18]. We address this gap by comparing both wearable (smartwatch) and mobile (smartphone) devices in a real-life production setting.

Smart industrial wearable or mobile devices can be used in various application scenarios. In particular, assistance systems are a prominent area of application. For instance Zenker and Hobert [37] utilize wearables to provide notifications during quality control processes. Aromaa et al. [2] qualitatively investigate how technicians assess wearable devices in maintenance scenarios. More generally, most studies examine selected sub-tasks of a production process, e.g., maintenance or assembly. In contrast, we holistically study the monitoring of the full production process. To do so, we equip shop-floor workers with wearable and mobile smart devices that support their diverse and complex everyday work.

2.2 Monitoring of Machinery in the Smart Factory

Physical machinery is at the heart of the production process. To facilitate failure-free operations, it is necessary to closely monitor machines. Usually, machine monitoring is conducted in office settings where visualizations are based on large screens. Examples include the visual diagnostic tool for assembly lines presented by Xu et al. [36], the routine monitoring and troubleshooting system for industrial purposes of Zhou et al. [40], or the visual equipment condition monitoring system of Wu et al. [35]. For an in-depth summary of related visualizations for the smart factory we refer to the review of Zhou et al. [39].

Monitoring of industrial machinery has also been applied to wearable and mobile devices. A few commercial products have already been available for several years (c.f. [37]). However, research and scientific knowledge on the topic remain scarce. Early work like the technical system of Wanbin and Tse [34] employs mobile phones and personal digital assistants (PDAs) for machine health monitoring. Aehnelt and Urban [1] present an assistance system prototype for shop-floor workers consisting of wearables and stationary monitors. Moreover, Padovano et al. [21] empirically evaluate the usage of a smartphone application enabling real-time condition monitoring and show that factory performance metrics might be improved. Usability aspects, however, were not addressed in this work. The laboratory-based work of Czerniak et al. [7] also suggests that using smartwatches and tablets for tasks like remote machine monitoring can improve machine operation. Specifically, reaction times might improve when using these devices.

We contribute to this body of research by presenting a machine monitoring system for both mobile (i.e., smartphone) and wearable (i.e., smartwatch) devices. This allows us to uncover advantages and disadvantages of the different types of devices.

2.3 Usability of Industrial Mobile and Wearable Devices

The usage of industrial wearable and mobile devices is characterized by distinct usability patterns, which arise from their form factor and interaction capabilities [17]. Early work by Siegel and Bauer [28] assessed usability aspects of wearables in the aerospace industry. Along similar lines, Url et al. [31] investigate the usability of an AR HMD with technicians in the automotive industry. Both studies found the mobile and wearable devices to be useful for practitioners, especially during unknown tasks. Papp et al. [23] explore usability aspects of a wearable body support system (exoskeleton). The authors find that such wearables might be perceived as a means to mitigate personal weaknesses, thereby having an impact of the users self-esteem.

Further work presents valuable prototypes, concepts, or lab-based studies, but lacks industrial context [2]. For instance, Kong et al. [14] present design considerations for industrial wearable systems based on a survey. Bröring et al. [6] describe a user-centered design process of a smartwatch app for machinery monitoring. Similarly, Baldauf et al. [3] discuss usability aspects of smart industrial wearable and mobile devices based on prototypes and emphasize, among other things, the need for compact notifications and multimodal feedback. The latter has also been studied by Funk et al. [10], who find that combining haptic and visual feedback might be suitable for communicating errors at the workplace. Mach et al. [17] assess user experience for an industrial smartwatch application compared to conventional monitor-based information display through a laboratory experiment with students. Here, the smartwatch was perceived as more attractive. Similarly, Vernim and Reinhart [33] compare usage frequencies of smartwatch and tablet in a simulated assembly task. While the tablet was used more frequently, it should be noted that the two types of devices offered different kinds of assistance: detailed instructions shown on the tablet and a call for expert button on the smartwatch.

Overall, usability aspects pertaining to industrial wearable and mobile devices remain largely unsolved [37], especially in real-life industrial settings. We close this gap by evaluating the usability of smart wearable and mobile devices under real-life conditions. Specifically, we conducted a field-study within a running factory under productive operation.

3 RESEARCH QUESTIONS

Based on the analysis of prior work, we defined the following three main research questions for our research.

RQ1: How do manufacturing workers experience ubiquitous machinery monitoring apps in the field?

As outlined in the previous section, prior work on mobile and wearable devices in factory settings has predominantly focused on assistance in concrete assembly and maintenance tasks, often through AR. In contrast, we address the comprehensive application of continuous machinery monitoring by workers. Specifically, we aim at studying the user experience and, in particular, the usability of novel wearable and mobile apps in a real-life production environment. These insights from the field can inform the design of user-friendly industrial ubiquitous monitoring solutions.

RQ2: How do manufacturing workers assess the impact of ubiquitous machinery monitoring on their work performance?
Due to the lack of scientific knowledge on ubiquitous machinery monitoring apps for workers, we investigate the workers' perceived impact of such apps on their work performance. We are interested whether they perceive advantages or disadvantages on their usual work routines through such an additional system. Based on workers' qualitative feedback regarding their experiences under productive operation, we derive recommendations for designing and deploying respective solutions.

RQ3: Do manufacturing workers prefer smartphones or smartwatches for ubiquitous machinery monitoring?
In order to best integrate ubiquitous machinery monitoring into the everyday work routines of manufacturing workers, we study the suitability of both smartphones and smartwatches for this purpose. In contrast to previous work, which often explored device preferences for different factory applications under lab settings, we investigate the favored device type for ubiquitous machinery monitoring in a running factory environment through a comparative study design.

4 UBIQUITOUS MACHINERY MONITORING SYSTEM

To answer the research questions put forth in Section 3, we designed and realized a worker-oriented machinery monitoring system under real-world conditions. This section introduces the cornerstones of this system and its development. We outline our participatory design approach for worker-oriented apps, give an overview of the system architecture, and describe the functional app prototypes.

4.1 Co-Design Workshops

To create both useful and easy-to-use monitoring apps for manufacturing staff, we followed a worker-centered co-design approach. We started with two requirements workshops with six and five production employees, respectively, to study their overall attitudes towards such apps and investigate their needs during the daily work. Together with two researchers, the participants identified typical intervention scenarios, explored the workers' information requirements, and discussed devices considered suitable for their work at the shop-floor. Two quality managers provided additional ideas for app features such as collecting error causes and the workers' responses for subsequent process analysis and optimization. During the workshops, we made use of device templates for sketching tasks, low-fidelity wireframes, as well as photo-realistic mockups to illustrate potential variants.

Having gathered fundamental requirements and co-designed early app sketches, first functional app prototypes were created. These prototypes were presented and discussed in two focus groups with six production workers to iteratively improve the apps. The researchers distributed devices with the prototypes pre-installed and used a custom simulator software to send out sample notifications on machinery failures to these devices. In a scenario-based walkthrough, the workers experienced typical situations and, following a think-aloud approach, commented on potential improvements of the apps. Suggestions included increasing font sizes for relevant information to enhance readability and shortening menu lists for quicker selections. For a detailed description of our co-design approach and corresponding design decisions, we refer to [3].

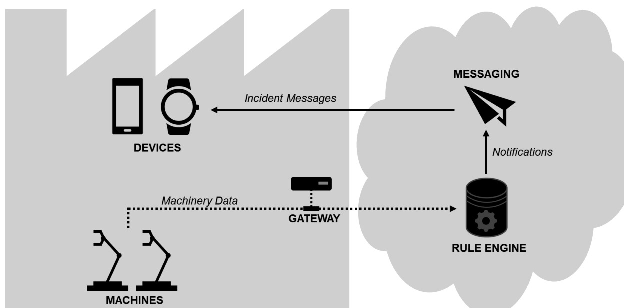

Fig. 2. High-level architecture overview of our monitoring system: Machinery data was continuously forwarded to a cloud-based monitoring system (right) via a gateway at the factory (left). An integrated rule engine checked for potential machine warnings and errors. For detected incidents, notifications were created and sent to the monitoring apps.

4.2 System Architecture

Figure 2 depicts the high-level architecture of our monitoring system. It builds upon Open Platform Communications Unified Architecture (OPC UA), a standard for accessing and exchanging machinery data in smart factory and Internet of Things (IoT) settings. OPC servers were installed on each machine to be monitored. They continuously forwarded real-time machinery data to a cloud-hosted management system for data storage, analysis and secure access. This system was realized using a flexible cloud-based toolkit for industrial IoT applications by *M&F Engineering AT*[1].

Based on the raw machinery data, rules for deriving high-level information and detecting failures and intervention scenarios were defined. In case an incident was detected, information such as the machine name, the timestamp, a short incident title, and the severity class (error/warning) were compiled into an incident description. Devices (or apps, respectively), which had registered for a given machine, were informed about a new incident through Firebase Cloud Messaging. Services to (un)register for incidents of a specific machine, retrieve incident lists, confirm an incident as fixed, etc., were offered through RESTful service interfaces secured via OAuth 2.0.

4.3 App Prototypes

Based on the workers' input during the co-design process, functional prototypes of two ubiquitous monitoring apps were implemented: one for a wearable device in the form of a smartwatch app (Figure 3), another one as a mobile app for smartphones (Figure 4). The wearable app was written for the *Wear OS* platform, an Android-based operating system utilized by smartwatch manufacturers such as Fossil or Samsung. The mobile app was built with *Angular* as a Progressive Web App, yet appearing as a locally installed mobile app for the users (e.g., through a respective app icon on the smartphone's home screen).

While the user interfaces were tailored to the device types according to established guidelines (e.g., [11]), the range of features was identical for the mobile and wearable app. At start-up, the users were asked to log in using a five-digit PIN (a unique part of their employee number) and a corresponding five-digit password. After a successful login, users selected from a list the machines

[1] https://m-f.ch

Fig. 3. The core screens of the monitoring app for smartwatches: login with employer ID, machine selection, main screen with time display and machines registered, an unsolved warning and error, detail screen of an error, and the selection of the measure taken (from left to right).

they wanted to receive notifications for. The main screens of the apps showed the current time (a requirement often mentioned during the co-design workshops) and a compact scrollable list of current (non-fixed) machine incidents. For each incident, the machine involved as well as the severity category were shown (errors in red, warnings in blue). The apps alerted to new incidents by sound (smartphone) and vibration (smartphone and smartwatch). In case the app was running in the background, new incidents were represented as OS push notifications (with sound and vibration) comprising the machine name, a name for the incident, and the severity category. Tapping either the list item, the OS notification, or the incident item opened a detail screen featuring a button for confirming the manual elimination of the incident. After pushing the button, the user was asked to select the reason for the incident as well as his/her response for solving the incident from two lists. These lists were compiled and condensed to five main items by experienced shift managers to avoid extensive scrolling, in particular on smartwatches.

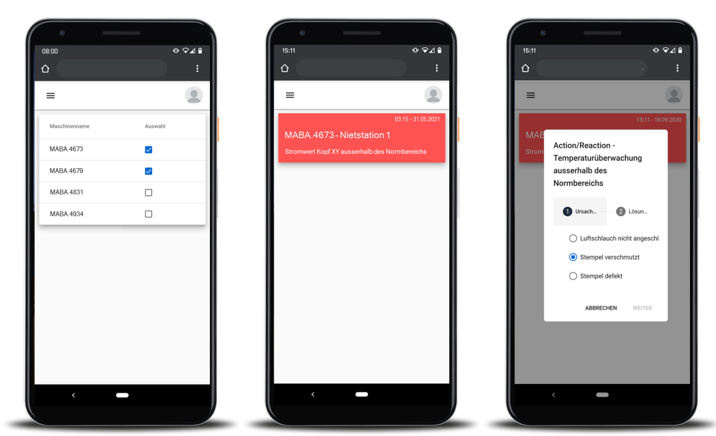

Fig. 4. Example screens of the monitoring app for smartphones: machine selection, list of current incidents, selecting reason and measure for an incident (from left to right).

5 METHOD

In this section, we describe our experiment's method in detail. We elaborate on the overall study context and design, the participants, devices used, as well as the data analysis.

5.1 Study Context

Utilizing the monitoring system and its apps, we designed and conducted a user study in the field at *Geberit Produktions AG* in Jona, Switzerland. This automated state-of-the-art production facility is part of the Geberit group, a globally active manufacturer of sanitary products. The company employs about 12,000 people worldwide and manufactures a variety of sanitary products. For the purpose of this work, we focused on the production of flush plates as depicted in Figure 5a, which are produced in Jona, Switzerland.

(a) Example of a flush plate. (b) Injection molding machine used in the study.

Fig. 5. Finished product (a) and industrial machine to produce the main components (b).

For the study, two machines at this Geberit production site were attached to the monitoring system. One of them is an injection molding machine (*ENGEL duo 350*, see Figure 5b), the other is a custom-made assembly machine. Both machines feature a warning lamp and an acoustic warning signal to indicate an incident as well as a small display to provide information on the incident. To generate corresponding push notifications on warnings and errors for these machines, 157 rules were defined. Table 1 shows four examples of rules. Corresponding notification texts (with the name of the machine concerned) were sent and displayed on the devices.

Machine	Rule	Severity Cat.	Notification Text
Molding	Welding point incorrect	Error	*Welding path out of norm range!*
Molding	Conveyor belt full	Warning	*Conveyor belt congested!*
Custom	Temperature of welding head low/high	Error	*Temperature of welding head out of norm range!*
Custom	Filling level of springs low	Warning	*Filling level of springs low!*

Table 1. Examples of rules and notifications for the molding injection machine and a custom-made assembly machine.

The workforce at the production site is composed of workers with different skills and job roles. However, since we focused on the production of a specific product type for the purpose of this study, the involved workers shared the same job profile and job roles. The main task of these employees was to ensure an uninterrupted production process. This included, among other things, refilling raw materials, cleaning machine parts, preparing machines for new product types, or monitoring machine uptime. The only difference among the workers was their assignment to different machines, so that each group of workers is responsible for a certain group of machines (e.g., an isle of machines). Section 5.4 provides more details regarding the participants recruited from the workforce.

5.2 Study Design

We applied a within-subjects design, where each worker involved, following a one-week pre-test phase, used both available app prototypes for two weeks (see Figure 6). The study was conducted in May and June 2021. We began the study with an introduction for the participants. During this event, we explained the idea of ubiquitous machinery monitoring and introduced the app prototypes in detail. Devices (i.e., smartphone and smartwatches) were randomly assigned to participants. Subsequently they were instructed on how to use the devices and the apps.

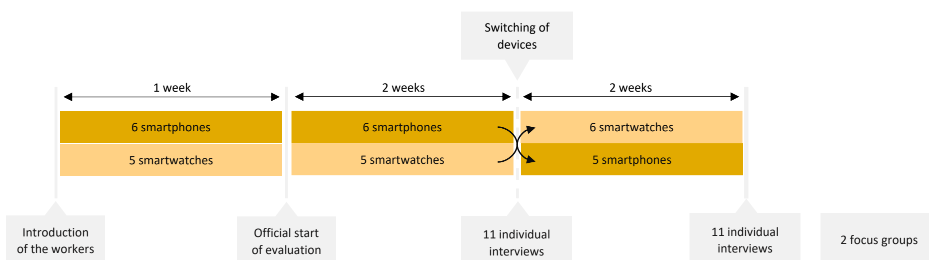

Fig. 6. After a one-week pre-test, our within-subjects design comprised two-weeks evaluation phases for each of the two app prototypes. Each evaluation phase was concluded by a set of individual interviews. Finally, two focus groups completed the study.

The one-week pre-test had two goals. First, we wanted the participating production workers to become familiar with the concept of ubiquitous machinery monitoring and corresponding novel processes such as picking up their device at the beginning of their shift, receiving messages about machinery incidents, confirming fixing an incident, or returning the devices at the end of the shift. Second, we aimed at identifying and solving potential technical problems, which might occur under production conditions, thereby impacting the workers' experience of the app prototypes during the evaluation phases of the study. Issues found and fixed during the pre-test included delayed receipts of notifications caused by vigorous battery management settings of the smart devices, an extension of session times to avoid frequent logins, or fine-tuning of thresholds for triggering certain predictive notifications.

After the pre-test, the participants used their device during their daily work at the shop-floor for two weeks. After this first evaluation phase, two researchers conducted semi-structured individual interviews with the participants regarding their experiences with the first app evaluated. Having completed an interview, the researchers handed out the second device (i.e., participants who had used a smartphone during the first phase received a smartwatch for the second phase and vice versa, see Figure 6) and instructed the participant about the second app. The participants then used their second device for another two weeks. During both evaluation phases, one researcher was present at the Geberit site for spontaneous support in case of a technical issue or a question regarding app usage. For each participant, the second evaluation phase was concluded by another individual interview. Finally, a few days later, the entire study was completed by focus groups with the involved production workers. Overall, we conducted 22 individual interviews and two focus groups throughout the study.

5.3 Individual Sessions and Focus Groups

Both individual sessions and focus groups were conducted by two researchers. While one acted as the main moderator, the other one was responsible for minuting and complementing follow-up questions. Both researchers took notes on the participants' responses and discussed and synthesized their notes into one protocol after each individual interview or focus group. In addition, all sessions were recorded and transcribed afterwards.

5.3.1 Individual Sessions. The individual sessions with the workers were conducted face to face and comprised a questionnaire and an interview with closed and open questions. The moderator followed an interview guide with several structured parts for collecting quantitative assessments.

The sessions comprised three main parts: the first assessed the usability of the app, the second collected overall feedback and experiences, and the third addressed the app's integration in the workers' daily work routines.

For evaluating the app usability, we used the well-established SUS questionnaire [5]. It contains 10 statements (e.g., I found the system unnecessarily complex, I thought the system was easy to use) and corresponding five-point Likert scales (from 1 - strongly disagree to 5 - strongly agree). Participants were asked to complete printouts of the questionnaire. The moderator accompanied this process and asked follow-up questions to collect participants' reasons.

In the second part, the moderator asked the participants to reflect on the following statements:

- "I would continue to use this app."
- "The app has a clear benefit for everyday work."
- "The app caused problems during my daily work."
- "The app gives me the feeling of being constantly informed about the production process."
- "With the app, I was able to react faster to machine incidents."
- "With the app, I could better concentrate on other work tasks."

For each statement the moderator asked the participants for their subjective assessment on a five-point Likert scale from strongly disagree (1) to strongly agree (5) as well as for their reasons.

The third part comprised questions on how the participants perceived the integration of the device/app into their work. This included receiving the device at the beginning of the worker's shift, its applicability and usage during the shift with its typical work tasks, handling the device during breaks, as well as returning the devices after the shift to a central storage place for charging. Again, quantitative assessments were collected on printed five-point Likert scales. Finally, the moderator asked for the participants' ideas for overall improvements of the apps and their potential concerns regarding the launch of such apps for ubiquitous machinery monitoring.

This three-part core of the guide was used for both individual interviews per participant, i.e., after the first and the second evaluation phase. However, the variant applied after the first evaluation phase was extended by an introductory part on the participant demographics such as age, sex, work experience, mobile device usage, and affinity to technology. The second variant (applied when participants had experienced both devices), concluded with the question, which device the participant would prefer for work, overall, and the corresponding reasons. One individual interview took about 30 minutes.

5.3.2 Focus Groups. For the focus groups, we composed participant groups of four to five production workers who had used both app prototypes before. Based upon the participants' responses during the individual interviews, we aimed at arranging groups of participants with varying impressions and opinions to stimulate discussion and gain further insights into the workers' requirements and experiences. One focus group took about 45 minutes.

Like the individual interviews, the focus groups were semi-structured along a set of guiding questions. These included the following:

- “Which device is more likely to be used further and why?”
- “What were the benefits of using the monitoring apps?”
- “Would you like to have such a monitoring solution launched at Geberit?”
- “What should be improved or added?”

5.4 Participants

Overall, 11 male production workers participated in the user study. They were aged between 25 and 52 years with a mean age of 38.1 years (SD=8.6). The participants’ work experience at their current workplace ranged from 3 to 26 years with a mean of 10.6 years (SD=8.7).

All participants were experienced mobile device users, using smartphones on a daily basis. Four participants owned and used a smartwatch, two additional participants used traditional wristwatches. Overall, the participants can be considered technology savvy: They showed major interest (4.6 of 5) and joy in exploring new technologies overall (4.3 of 5) and assessed their perceived overload through novel technologies as rather low (1.9 of 5).

5.5 Devices

Six smartphones and six smartwatches were prepared for the participants. As smartphones we used *Motorola Moto E6 Plus*. These devices run Android 10 and feature a 6.10" display with 1560x720 pixels. As smartwatches we used *Fossil Sport* watches. They are powered by Wear OS 2.23 and their circular 1.19" display consists of 390x390 pixels. The smartwatches were configured over Bluetooth via the corresponding Android app *Wear OS*. However, during the study, the smartwatches and the monitoring app operated in stand-alone mode without any smartphone app connected. All devices were connected to a shop-floor-wide WiFi for accessing the Web and receiving notifications.

5.6 Ethical Considerations

Since our research involved employees and touched upon sensitive topics such as work attitude, satisfaction, and performance, ethical aspects were given special consideration. During the introductory event, the participants were instructed about the purpose of the study, i.e., the investigation of monitoring apps from the workers’ perspective. It was emphasized that participating in the study and using the available app prototypes was voluntary. Similarly, withdrawing from the study and the interviews was allowed at any time without the need for providing a reason. The participants’ consent and permission to collect usage and logging data from the apps, to record interviews, and to process these data was obtained.

The researchers assured that all data collected would be used only for the study purpose communicated. This excluded analyses and conclusions regarding the participants’ individual work performances. For example, response times after receiving a notification or no-response ratios after receiving a notification were not calculated. Furthermore, all data collected was pseudonymized before the analysis, i.e., personal information that might identify individual workers were removed and participants’ names replaced with artificial identifiers. All data sheets were stored at a secure cloud service only accessible for the involved research team, not any management staff.

5.7 Data Analysis

Our research had a qualitative focus. We conducted a reflexive thematic analysis [4] of the participants’ responses in both individual interviews and focus groups with the goal to identify and interpret themes and patterns. Having familiarized themselves with the data by reading and rereading the transcripts, two researchers coded the responses in a collaborative manner for sense-checking and exploring multiple interpretations of the data. Instead of using a structured codebook, the researchers iteratively created a list of codes (starting from prior related work on industrial mobile and wearable apps). Following an inductive approach, themes were then created from the codes collaboratively. Overall, we did not aim at achieving consensus among the coders but rather exploring richer interpretations of the data. We selected verbatim quotations (translated to English by the researchers) to illustrate themes relevant for answering the research questions.

For the participants' quantitative responses (i.e., their agreement to the various statements), descriptive and inductive analyses in SPSS were conducted. Since we applied five-point Likert scales, we treated each variable as an ordinal approximation of a continuous variable (cf. [13]). We ran dependent t-tests in SPSS to check for significant effects of the device type.

6 RESULTS

This section presents the results of our field study in detail. We elaborate on participants' responses in individual interviews and focus groups.

6.1 Quantitative Questionnaires

In the following, we report on the participants' quantitative responses during the individual interviews.

6.1.1 SUS Evaluation. Both prototypes received very high SUS scores. The mean rating for the smartphone app was 95.2 (SD=4.3) with a minimum rating of 85 and a maximum rating of 100. The mean rating for the smartwatch app was slightly higher with 95.9 points (SD=5.4), again with a minimum rating of 85 and maximum rating of 100.

6.1.2 Assessments of Acceptance, Implications, and Applicability. Figure 7 depicts the participants' mean assessments of different UX-related statements regarding the acceptance, implications, and applicability of the devices and prototypes. We observed statistically significant differences for one statement relating to the devices' applicability during typical work tasks ("Using the device worked well"): The smartphone received a mean score of 2.82 (SD=1.89), the smartwatch a mean score of 4.82 (SD=0.60), t(10)=-3.028; p=.013.

All other statements were rated similarly for the two device types without statistically significant differences. The statement "I would continue to use the app" was answered positively overall, with a mean of 3.91 for the smartphone app (SD=1.30) and a mean of 3.82 (SD=0.98) for the smartwatch app, t(10)=.247, p=.810. Similarly, the statement "The app has a clear benefit for my job" received a

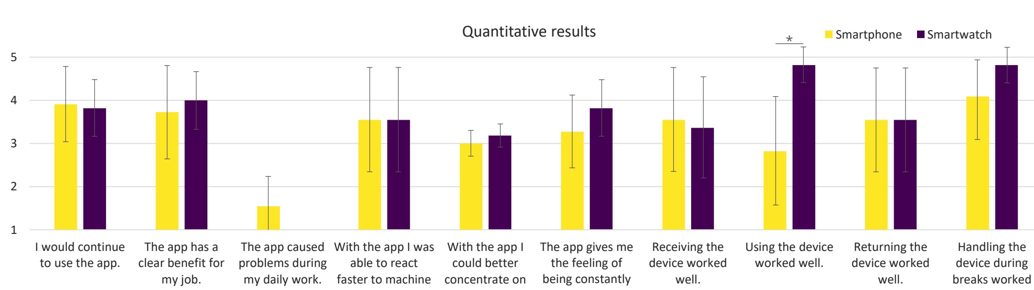

Fig. 7. The participants' mean agreement to statements during the 22 individual interviews on Likert scales from 1 (strongly disagree) to 5 (strongly agree). Error bars represent 95% confidence intervals.

mean score of 3.72 (SD=1.62) for the smartphone app, and a mean score of 4.00 (SD=1.00) for the smartwatch app, t(10)=-.521, p=.614. The statement "The app caused problems during my daily work" was denied by the participants, overall. The scores were 1.55 (SD=1.04) for the smartphone app and 1.00 (SD=0.00) for the smartwatch, t(10)=1.1.747, p=.111.

The participants confirmed the statement "With the app, I was able to react faster to machine incidents" with a mean rating of 3.55 (SD=0.69) for the smartphone, with 3.55 for smartwatch (SD=0.93), t(10)=.0, p=1.0. Regarding the statement "With the app I could better concentrate on my actual tasks", participants were rather neutral and provided a mean rating 3.00 (SD=0.45) for the smartphone and of 3.18 (SD=0.40) for the smartwatch, t(10)=-.803, p=.441. The statement "The app gives me the feeling of being constantly informed about the production process" was rated with a mean score of 3.27 (SD=1.27) for the smartphone, and 3.82 (SD=0.98) for the smartwatch, t(10)=-1.067, p=.311.

Finally, participants were asked to assess the integration of the devices and the corresponding apps into their daily work routines. Besides the significant differences for the statement on the devices' applicability, the remaining statements were rated similarly. The reception of the devices was rated with 3.54 (SD=1.81) for the smartphones and 3.36 (SD=1.74) for the smartwatches, t(10)=.289, p=.779. Returning the devices was rated equally for both the smartphones and the smartwatches with 3.55 (SD=1.81), t(10)=.0, p=1.0. Finally, the participants rated the handling of the smartphone app during breaks with 4.09 (SD=1.38), of the smartwatch app with 4.81 (SD=0.60), t(10)=-1.491, p=.167.

6.1.3 Overall Device Preference. When asked for their overall preference for one of the two devices, we found strong consistency among the participants: In direct comparison, all of the 11 professionals favored the smartwatch over the smartphone for continuous machinery monitoring during their daily work.

6.2 Individual Interviews

In the following, we report on the participants' qualitative responses, structured along the main sections of the individual interviews.

6.2.1 SUS Evaluation. In their explanations of the high SUS scores, the participants appreciated the simplicity of the apps. One participant referred to his prior hesitations, when being introduced to the project: *"I had concerns, that it's going to be complicated with these devices, but it was great."* (P7). Furthermore, they valued the compact presentation of the notifications and described the incident details shown as sufficient. In addition, four participants positively mentioned the colors of the severity categories which were equivalent to the color coding of incident alerts at the machines' displays and lamps.

6.2.2 Acceptance of Monitoring Apps. When asked for the benefits of the apps for their job, all participants emphasized the advantage of the monitoring apps for situations, *"when the machines are out of the worker's sight or cannot be heard"* (P4). Statements included *"It's a good overview of the machines, but the apps are most helpful for machines that are not visible, for example ones in another sector"* (P2), *"When I am not at the shop-floor, I am working between the machines, or I am having a break, then the apps are useful. Otherwise we automatically observe the lamps at the machines."* (P4), and *"It's useful when I'm just busy somewhere else and don't see the machine."* (P5). Two participants referred to their role as floaters who have additional tasks such as cleaning and delivering (intermediate) products. They considered the apps, which drew their attention to defective machines, as particularly useful.

Relevant Information. Several participants considered the value of the monitoring apps dependent on the information given by the notifications. While the occurrence of a production-critical error is widely visible at the shop-floor through the lamps mounted at the machines, relevant production parameters can not be seen from a distance. An example mentioned by six participants was the cycle time, which describes how long it takes a machine to produce one unit of a given product: *"Problems regarding the cycle time can only be detected by workers with great experience, they are able to hear from the sound of the machine but only if they are close to the machine, of course."* (P1). Similarly, five participants considered predictions such as notifications on an upcoming shortage of material helpful. For instance, if there are not enough rubber seals in the supply container at the machine, the production process cannot continue.

Specific work situations. Two participants noticed additional benefits of the apps' notifications related to special work situations. P9 referred to situations where workers give special attention to one specific machine, either since this machine is pivotal for the current production or it is currently under repair. Other machines might be visible, however, are ignored during intensive repair activities. *"Sometimes I'm completely focused on one machine, then the app is very helpful to keep an eye on the other machines"*. P1 indicated extensive inspection rounds, particularly in production halls with widely distributed machinery and considered such apps useful in cases where he does not feel physically fit: *"I need to walk a lot between the machines. When I had a bad night and am tired, I might walk less during my shift but still keep an overview of the machines."*

Support for novices. Particular value of the monitoring apps was seen for new colleagues with less work experience. While the experienced workers were very proficient in overseeing the existing alert lamps, even during other work and maintenance tasks, the participants considered *"[the apps] very useful , so the beginners can rely on the app."* (P4). In particular, relevant indicators such as cycle times might be easily checked by novices via such apps, according to the participants.

Lengthy data input. While the participants did not see major problems with the apps for their work tasks, some of them brought up capturing the incident reason and the worker's response as a lengthy procedure. P6 considered it *"[...] an additional effort to confirm and select the reason"*, P7 complained that *"[...] answering these questions takes too long when there are a lot of error messages"* and suggested that this function is removed.

Negligible privacy concerns. When asked for their thoughts on privacy and potential fears of being tracked by the company through the devices (e.g., reaction times, distances covered, vital parameters), no participant expressed severe concerns. While several participants reasoned this with their trust in the company, others referred to their work performance and said they *"[...] have nothing to hide, because we do a good job."* (P10). Only one participant expressed the slight feeling of additional stress since he assumed that *"the boss receives these messages too"* (P5) and thus felt that he needed to tackle incidents immediately. Three participants described an uncomfortable feeling when starting to use the study devices since it is not allowed to use private smartphones at the shop-floor and they were afraid to appear idle. For example, P1 described *"[...] odd looks by colleagues because they thought I'm playing around with my mobile"*.

Differences to full implementation. In their verbal comments regarding a potential further usage of the apps, a few participants mentioned their concerns regarding a fully productive realization for the entire machine park. Examples include *"The notifications are helpful, but I'm afraid of a lot of notifications when additional machines are integrated"* (P5) and *"When more machines will be connected, we will receive too many error messages."* (P7).

6.2.3 Implications on Work Routines. Four participants described they were able to react faster. These participants provided two examples. One involved a supply mechanism of metal springs, which needs to be filled regularly and checked for potential jamming. The other pertained to a mismatch of the targeted rate of production and the actual rate of production as expressed through the cycle time of a machine. Four others felt no improvements regarding reaction times through the apps. One participant referred back to a prior remark on the lengthy confirmation process and assumed shorter reaction times when the apps were only presenting information on incidents and would not require users to select incident reasons and actions.

Regarding an improved concentration on actual work tasks, four participants explained their low ratings by their lack of trust in the apps, e.g., *"I think I could really better concentrate on other tasks if I know for sure that the app works perfectly and I can trust it."* (P2). One participant illustrated his positive rating regarding the constant information through the monitoring apps as follows: *"It's a good feeling to know that the machines are working properly. It's one less thing I have to worry about."* (P1).

6.2.4 Integration into Everyday Work. The question regarding the devices' applicability during typical work tasks provided diverse answers. Eight participants complained that they carried around two (private and study) or even three (private, business, and study) mobile phones in their pockets and that the devices started to become heavy during the working day. There were no related remarks regarding the smartwatch. Concerns regarding damages to the devices were scarce. Only two participants expressed minor fears of breaking the smartwatches at a machine during repair tasks and described that they usually take off their private wristwatches for work. One participant mentioned slight concerns regarding safety at work with the smartwatch: *"Once I was afraid to get stuck at a machine with the wristband."* (P10).

Storage location for devices. Several participants remarked that the devices should be provided at a location regularly passed by the workers on their way to the shop-floor, such as the respective shift leader's desk. In the study setup, all devices were stored and charged at a central location in one of the production halls. In consequence, some workers had to make a detour to get their device. In a few cases, participants forgot to pick up their device at the beginning of the shift. *"When I arrive in the morning, I'm immediately stressed when I see that some machines are not working properly"*, as P1 explained the forgetting of the device. P7 argued for a storage location very close to the actual working place, too, because *"it would be easier to check whether the device is charging correctly"*. Since one one device was lost during the study and this worker then used the one of a colleague from another shift, he came up with the idea of lockable containers for each worker to securely store the device. The participants' feedback regarding issues when returning devices mainly addressed the suitable location of storing the devices, e.g., *"Since I wasn't reminded, I sometimes even forgot to bring back the watch. The shift leader's desk would be good solution."* (P1).

Non-working times. With regard to handling the devices and apps during breaks, five participants suggested a "pause" button to silence notifications for the duration of the break (typically 30 minutes). Most of the participants took the devices to the break room and did not explicitly express that they felt disturbed by the notifications, yet several mentioned that *"... [the notifications] could become annoying when more machines are connected"* (P7). Only one participant left the smartphone at the storage location to avoid being disturbed during his break. Four participants noticed advantages of the monitoring apps during their breaks. They referred to situations when the substitute worker is less experienced or there is no substitute worker at all. Thus, they appreciated staying informed about the production process also during breaks. *"In some cases, for example at weekends, I don't have a deputy colleague. Then it's very useful to also have an eye on the machines during breaks."* (P4).

6.2.5 Overall Device Preferences. While all participants favored the smartwatch over the smartphone in direct comparison, one participant saw *"... more potential in the smartphone solution"* (P8). He expected more features such as documenting error messages at machines (utilizing the built-in camera) to be added for the mobile app. For the smartwatch app, he was concerned about the worsened usability when more features were integrated.

The main advantages of the smartwatch mentioned by the participants included the direct and quick interaction with the wearable device. Respective statements included *"The watch is much better, I recognize messages earlier"* (P3) and *"It's much easier, because you watch the smartwatch automatically."* (P7). The necessity of taking the smartphone out of a trouser pocket to interact with the mobile app and read about machine incidents turned out to be a core drawback at the shop-floor. The vibration alert of the smartwatch, on the other hand, was highly appreciated by the participants: *"I immediately noticed incoming notifications through the vibration and saw the information at my wrist."* (P5).

Two participants without prior smartwatch experience mentioned that it took them more time to deal with the wearable device and the respective app than with the smartphone. Still, both considered this learning phase short and reported to have quickly developed routines. Several participants described great wearing comfort regarding the smartwatch, e.g., *"After three days I sometimes didn't even notice that I was wearing the smartwatch."* (P1).

Most of the participants did not report on major problems when interacting with the smartwatch. Only two participants suggested using smartwatch models with a larger display size to facilitate controlling the app and reading about incidents. Two other participants mentioned increased sweating at the wrist due to the plastic wristband of the smartwatch model used and suggested a wristband made of fabric or leather. Beyond the devices investigated, one participant suggested a vibrational wristband to draw the workers' attention to the alert lamps of the machines in case of incidents.

Regarding the smartphone and carrying it in pockets, most participants complained about not (immediately) noticing the incident messages in several cases. Vibrations were not felt due to the rugged work pants or movements during physical tasks, audio alerts not due to ambient noises in the factory or due to earplugs worn by some workers. In addition, several participants had concerns about carrying the smartphone in their pockets, since they contain hard and/or sharp objects such as keys, knives, and other tools.

6.3 Focus Groups

During the focus groups, the participants emphasized several core aspects from the individual interviews. Particular topics that led to intense discussions among the participants included how to avoid the workers' potential overload by notifications and how to efficiently deal with a large number of notifications. In the following, we introduce the main themes that emerged during the focus groups and present participants' pivotal responses.

6.3.1 Notification Overload. Several participants repeated their concerns regarding an unmanageable number of notifications in case additional or even all machines were connected to the ubiquitous machine monitoring system. During the participants' discussion, the manual confirmation of the notifications turned out to be the main problem. Statements included *"If it's only one machine, that's no problem, but confirming the notifications for a lot of machines will take too much time"* (P7) and *"Now I need to confirm twice: at the machine and the device."* (P8). Three participants highlighted this drawback in particular for the smartphone, which the workers had to take out of their pockets to confirm the failure after they had confirmed it at the machine. The participants agreed that no manual confirmation of an incident should be required at all: *"If I remedy a defect, the message on the device should vanish"* (P1). Similarly, messages for incidents which might dissolve over time (e.g., notification of upcoming end of production), should also disappear from the devices when the incident is over, as one participant recommended.

6.3.2 Suspending Notifications. The idea of a "pause" button for silencing notifications was pursued further during the focus groups. Originally supposed for work breaks, several participants appreciated a related function also for cases when a machine is being adjusted, it stands still, or a relevant sensor is broken. Then, the rule-based system continuously generates notifications since thresholds are either undershot or exceeded. In particular, the participants complained about a lot of messages during the configuration of machines: *"This is cumbersome, there are lots of notifications until a machine works properly."* (P6). P1 concluded that *"there should be as few errors in the production as possible, so that such an app makes sense"*. Two participants referred to mobile messengers, which allow to temporarily mute senders. Similarly, in the monitoring apps machines could be muted for predefined periods of time (e.g., 30 or 60 minutes), so that workers in, for example, stressful situations, do not forget to unmute.

6.3.3 Data Input. While most of the participants considered the selection of error causes and workers' actions too time-consuming and inapplicable in a productive factory environment, one participant identified an advantage of this feature. If a worker cannot repair a malfunction himself, he or she currently needs to inform the shift manager or the maintenance team by phone. Instead, the participant suggested, the notification could be forwarded to them when the corresponding action is selected within the app. *"Additionally, feedback by the maintenance team could be shown in the app. For example, that they are on the move or when they will arrive"* (P2). Another participant expanded on that thought and suggested adding a button for calling the maintenance team next to the button for manually confirming an incident.

6.3.4 Feature Ideas. In addition, the participants created and discussed several ideas for additional features and overall improvements. For the main screen showing the current time, they suggested adding the current date and day of the week. Additional features requested included custom timers for certain tasks. For instance, some paints need to be mixed and some machine parts need to be cleaned at regular intervals. Two participants envisioned respective time-based reminders within the app. Regarding the notifications about the cycle times, two participants found information about the slowest components within a production line useful. One participant suggested utilizing the smartphone camera for documenting repair processes and sharing the photos within the apps, *"[...] to show beginners how the machine needs to look after troubleshooting."* (P8).

6.3.5 Shared vs. Personal Devices. Finally, the participants agreed that devices should not be shared but each worker should have a personal device. They not only argued this point for hygienic reasons but also thought that workers might feel more responsible for a personal device and thus take better care of them. Furthermore, workers could store a personal device at a location they deem suitable for their work routines, in contrast to a shared device, which needs to be picked up at a central place when the shift starts.

7 DISCUSSION

In this section, we refer back to our research questions and discuss the results. Furthermore, we outline limitations of the study.

7.1 RQ1: How do manufacturing workers experience ubiquitous machinery monitoring apps in the field?

Overall, the workers provided positive feedback on the monitoring apps. Both prototypes received very high SUS scores and no major usage problems were reported when using the apps throughout the study. We ascribe this to the participatory design of the monitoring apps, which led to a compact presentation of the relevant information, a well-usable navigation on both smartwatch and smartphone, the stripped-down functionality, as well as the use of conventions such as established color codes. While most workers considered the apps useful and stated they would continue using them, we found major concerns regarding the workers' overload by machine notifications. A volume of notifications might be caused by a misconfiguration of a production line or a specific machine. Furthermore, additional machines integrated in a fully productive setup will increase the number of notifications. For such cases, a mute function is a relevant feature for workers to cope with a large number of messages, remain in control of the system, and avoid distraction.

Privacy concerns regarding the mobile and wearable devices did not play a major role, neither were monitoring apps rejected due to these. However, following the participants' statements on trust, this might be dependent on the workers' attitude towards the employer.

7.2 RQ2: How do manufacturing workers assess the impact of ubiquitous machinery monitoring on their work performance?

Workers perceived the impact of the apps on their work performance very differently. While the related quantitative responses were often positive (e.g., "With the app I was able to react faster to machine failures"), we found that several parameters affect the perceived impact of the apps and their usefulness. First, workers are highly trained to observe machines' alert lamps, which, in the case of the factory studied, indicate failures and warnings through red and blue lights. Thus, in open and overseeable production areas, workers usually do not benefit from monitoring apps when they simply duplicate these incident notifications. An exception are cases where workers are busy with lengthy repair tasks drawing their attention to a specific machine or when they have additional tasks requiring them to temporarily leave the shop-floor. Second, a particular benefit is provided by notifications beyond information on typical machine incidents. Examples include indicators, which are not easily apparent (e.g., cycle times) and predictions (e.g., indications of containers soon to be filled).

In addition, we found interactive functions beyond the actual notification mechanism, in our case particularly the selection of reasons of an incident and the workers' responses, to impede workers. Although the lists of available reasons and actions were shortened by experts, context-aware (i.e. tailored to the respective incident), and well usable on both devices, the workers felt this feature too time-consuming and distracting from their actual work. While prior work has considered capturing workers' knowledge through devices promising (e.g., [3]), our insights from the field show the arising drawbacks for workers in their daily work. Specifically, collecting responses in an unobtrusive way that is accepted by workers remains challenging. Potential remedies might include context-aware mechanisms to prompt workers for data input only in opportune moments (see e.g., [9]) as well as gamification and nudging strategies (see e.g., [8, 32]). While gamification approaches have been applied successfully in various applications, respective experiments (in particular field studies) in the manufacturing domain are scarce. Since workers complained about the additional effort of data input during their work activities, we see more potential in prompting for input in a context-aware manner. For example, suitable, less stressful moments might be detected by considering information on the overall production process or even workers' physiological parameters collected through smartwatch sensors (cf. [19]).

7.3 RQ3: Do manufacturing workers prefer smartphones or smartwatches for ubiquitous machinery monitoring?

While most quantitative assessments of the two device types and the corresponding apps showed comparable results and the usability scores given were high for both, asking for the preference yielded a clear result: Workers unambiguously favored the smartwatch and the corresponding app. They appreciated the well noticeable vibration alert, having crucial information present at the wrist as well as the smartwatches' overall wearing comfort. While only four of the eleven participants had prior experience with smartwatches, introductions in using the smartwatch and the monitoring app were brief and no major usage problems were reported during the field trial.

Experiences from the field study indicate that smartphones, typically carried in trouser pockets during work in the factory, are not suited for monitoring tasks that require quick reactions. Workers often miss both vibration and audio alerts and thus notice the respective messages only when they take the smartphone out of the pocket. While data glasses have been studied for AR-based assembly tasks [2, 12, 20], they have been rejected by workers for long-running monitoring task [3].

To the best of our knowledge, our comprehensive field study demonstrated the applicability and benefits of smartwatch apps over mobile devices over several weeks under real-world conditions for the first time. Prior work on smartwatches for Industry 4.0 settings (e.g., [1]) either expected advantages and/or conducted related studies with laymen or in lab environments (e.g., [17]).

7.4 Limitations

We conducted a comprehensive user study in a real-life production environment for five weeks (including one week of functional testing of the prototypes) and gained novel insights into the user experience of worker-oriented monitoring apps. Workers' trust in such a monitoring solution was found to be an essential factor for effective assistance and potential efficiency increases. Still, the investigation of this trust building as well as long-term effects of such apps on the workers themselves and their work routines need additional research.

In our study, participants covered a broad age range. However, according to their self-assessment regarding interest and joy in exploring new technologies, the group contained several tech-savvy workers. The very high SUS scores and positive responses during the introductions, the study, and the interviews suggest similar feedback for a very broad user group. Still, a complementary study could focus on potential challenges for and concerns of a less tech-savvy participants sample. Regarding the explorative nature and qualitative focus of our study with a limited number of participants, a future study might also include a larger sample size to validate our quantitative findings.

We found indications that the workers' attitudes towards to the employer and the overall corporate culture might have had an impact on the workers' acceptance and experience of machinery monitoring apps. Since our field study was conducted within one company's production facilities, we are not able to draw comparisons between companies. Respective in-depth investigations and conclusions are therefore not possible and remain subject to future work.

8 RECOMMENDATIONS FOR UBIQUITOUS MACHINERY MONITORING

Based on the results of our field study, we derive a set of recommendations for designing and deploying ubiquitous machinery monitoring apps for factory workers.

Consider wearable monitoring apps. Despite their limited display size, smartwatches are preferred over smartphones by factory workers for continuous monitoring of machinery, when directly compared. Advantages include the vibration alert directly on the wrist, quick access to the information as well as the devices' wearing comfort. The display size of mass-market smartwatches is sufficient for presenting relevant information on machine incidents. In contrast, incident notifications on smartphones, which are usually carried by workers in trouser pockets, tend to be missed.

Reduce explicit app interactions. Machinery monitoring apps should focus on presenting information in a compact form while avoiding explicit interactions as much as possible. Collecting additional information through manual entry or selection is time-consuming and distracting for workers and should be implemented cautiously. Even simple manual confirmations of reported incidents can become cumbersome in case of a vast number of active notifications. Hence, incident messages on devices should automatically vanish as soon as the respective machine has been repaired.

Provide personal devices. Each worker should be assigned an individual device instead of sharing devices between workers in different shifts. This allows workers to customize the device according to personal preferences and to best integrate the devices into their routines. Examples include choosing a textile wristband for a smartwatch over a plastic one to avoid sweating or deciding for an optimal storage location to easily pick up and return the device for charging.

Choose messages consciously. In an open production area where all machines are well-visible from various workplaces, workers are usually aware of current failures due to warning lights at the machines. However, they benefit from notifications regarding preventive interventions (e.g., "container 90% full", "fill level below 10%"). In contrast, in wide or multistory production facilities (and for workers often on the move), device notifications are useful for making machine messages remotely available. Furthermore, information that might not easily be apparent at machines, e.g., divergent cycle times, job ends, or job changes, are worth considering.

Prevent message overload. A myriad of simultaneously active machine notifications on a mobile or wearable device does not support workers but is irritating and distracting. Besides the careful selection of rules (e.g., to inform only about frequent or serious incidents), system features should be considered to avoid worker overload. Notifications might include time-outs to remove old messages (e.g., created during the previous shift) without the user's explicit confirmation. In addition, workers should retain control over the notification mechanism, e.g., through a mute function to temporarily silence or block notifications (entirely or for specific machines).

Ensure production quality. As a further means to prevent a message overload and to make incident notifications effective, a certain level of production quality is necessary. For example, many threshold-based rules can be triggered during the setup and adjustment of single machines and the entire production line as long as the manufacturing process has not reached a production-ready quality. Besides a device-side mute function mentioned above, a mechanism for centrally disabling notifications for these phases may be considered.

9 CONCLUSION AND OUTLOOK

In this paper, we presented an extensive explorative field study on ubiquitous machinery monitoring with manufacturing workers. 11 professionals evaluated monitoring apps on smartwatches and smartphones for four weeks and reported on their experiences and impressions in 22 individual interviews and two focus groups. To the best of our knowledge, this evaluation represents the first multi-week field study on mobile and wearable machinery monitoring in a running factory.

Overall, the concept of ubiquitous machinery monitoring and the apps received positive feedback. Still, we found concerns regarding tedious data entry and potential overload by machine notifications. The perceived benefit was found to be dependent on the production setup and the message types. Finally, smartwatches turned out to be clearly favored over smartphones for ubiquitous machinery monitoring tasks in manufacturing settings. Based on these results, we derived a set of recommendations for designing and deploying ubiquitous machinery monitoring apps for factory workers.

This study deliberately focused on the workers' user experience of the respective monitoring apps and subjective impressions. To complement this work, future research could study the impact of ubiquitous machinery monitoring approaches on relevant performance indicators. Moreover, from an HCI perspective, collecting contextual information on the production state and workers' actions through wearable devices seems worth investigating in depth.

REFERENCES

[1] Mario Aehnelt and Bodo Urban. 2014. Follow-Me: Smartwatch Assistance on the Shop Floor. In *HCI in Business*, Fiona Fui-Hoon Nah (Ed.). Springer International Publishing, Cham, 279–287.

[2] Susanna Aromaa, Iina Aaltonen, Eija Kaasinen, Joona Elo, and Ilari Parkkinen. 2016. Use of wearable and augmented reality technologies in industrial maintenance work. In *AcademicMindtrek 2016 - Proceedings of the 20th International Academic Mindtrek Conference*. 235–242. https://doi.org/10.1145/2994310.2994321

[3] Matthias Baldauf, Sebastian Müller, Arne Seeliger, Tobias Küng, Andreas Michel, and Werner Züllig. 2021. Human Interventions in the Smart Factory – A Case Study on Co-Designing Mobile and Wearable Monitoring Systems with Manufacturing Staff. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3411763.3451774

[4] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* American Psychological Association, 57–71. https://doi.org/10.1037/13620-004

[5] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (11 1995).

[6] Andre Bröring, Arno Fast, Sebastian Büttner, Mario Heinz, and Carsten Röcker. 2019. Smartwatches zur Unterstützung von Produktionsmitarbeitern. (2019). https://doi.org/10.18420/MUC2019-UP-0324

[7] Julia N. Czerniak, Nikolas Schierhorst, Christopher Brandl, Alexander Mertens, and Verena Nitsch. 2020. Smart digital assistance devices for the support of machine operation processes at future production workplaces. *Advances in Intelligent Systems and Computing* 1217 AISC (2020), 491–497. https://doi.org/10.1007/978-3-030-51828-8_64

[8] Rodrigo De Oliveira, Mauro Cherubini, and Nuria Oliver. 2010. MoviPill: Improving medication compliance for elders using a mobile persuasive social game. *Proceedings of the 2010 ACM Conference on Ubiquitous Computing* (2010), 251–260. https://doi.org/10.1145/1864349.1864371

[9] Joel E. Fischer, Chris Greenhalgh, and Steve Benford. 2011. Investigating episodes of mobile phone activity as indicators of opportune moments to deliver notifications. *13th International Conference on Human-Computer Interaction with Mobile Devices and Services* (2011), 181–190. https://doi.org/10.1145/2037373.2037402

[10] Markus Funk, Juana Heusler, Elif Akcay, Klaus Weiland, and Albrecht Schmidt. 2016. Haptic, auditory, or visual? Towards optimal error feedback at manual assembly workplaces. *ACM International Conference Proceeding Series* 29-June-20 (2016), 1–6. https://doi.org/10.1145/2910674.2910683

[11] Google. 2021. Principles of Wear OS development. https://developer.android.com/training/wearables/principles. Accessed: 2022-02-01.

[12] Jérôme Jetter, Jörgen Eimecke, and Alexandra Rese. 2018. Augmented reality tools for industrial applications: What are potential key performance indicators and who benefits? *Computers in Human Behavior* 87, February (2018), 18–33. https://doi.org/10.1016/j.chb.2018.04.054

[13] David Richard Johnson and James C. Creech. 1983. Ordinal Measures in Multiple Indicator Models: A Simulation Study of Categorization Error. *American Sociological Review* 48, 3 (June 1983), 398. https://doi.org/10.2307/2095231

[14] Xiang T.R. Kong, Hao Luo, George Q. Huang, and Xuan Yang. 2019. Industrial wearable system: the human-centric empowering technology in Industry 4.0. *Journal of Intelligent Manufacturing* 30, 8 (2019), 2853–2869. https://doi.org/10.1007/s10845-018-1416-9

[15] Christian Krupitzer, Sebastian Müller, Veronika Lesch, Marwin Züfle, Janick Edinger, Alexander Lemken, Dominik Schäfer, Samuel Kounev, and Christian Becker. 2020. A Survey on Human Machine Interaction in Industry 4.0. (2020), 45 pages. http://arxiv.org/abs/2002.01025

[16] Francesco Longo, Letizia Nicoletti, and Antonio Padovano. 2017. Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. *Computers and Industrial Engineering* 113 (2017), 144–159. https://doi.org/10.1016/j.cie.2017.09.016

[17] Sebastian Mach, Almut Kastrau, and Franziska Schmalfuß. 2018. Information at hand – Using wearable devices to display task information in the context of industry 4.0. *Communications in Computer and Information Science* 850, June 2019 (2018), 93–100. https://doi.org/10.1007/978-3-319-92270-6_13

[18] Benedikt G Mark, Erwin Rauch, and Dominik T Matt. 2021. Worker assistance systems in manufacturing : A review of the state of the art and future directions. *Journal of Manufacturing Systems* 59, March (2021), 228–250. https://doi.org/10.1016/j.jmsy.2021.02.017

[19] Sebastian Müller, Matthias Baldauf, Andreas Michel, and Peter Fröhlich. 2021. Advanced Ubiquitous Monitoring Services for Workers in Automated Production Environments. In *Proceedings of Automation Experience at the Workplace.* CEUR-WS. http://ceur-ws.org/Vol-2905/paper10.pdf

[20] Andrew. Y.C. Nee, Soh . K. Ong, George. Chryssolouris, and Dimitris. Mourtzis. 2012. Augmented reality applications in design and manufacturing. *CIRP Annals - Manufacturing Technology* 61, 2 (2012), 657–679. https://doi.org/10.1016/j.cirp.2012.05.010

[21] Antonio Padovano, Francesco Longo, Letizia Nicoletti, and Giovanni Mirabelli. 2018. A Digital Twin based Service Oriented Application for a 4.0 Knowledge Navigation in the Smart Factory. *IFAC-Papers OnLine* 51, 11 (2018), 631–636. https://doi.org/10.1016/j.ifacol.2018.08.389

[22] Volker Paelke, Carsten Röcker, Nils Koch, Holger Flatt, and Sebastian Büttner. 2015. User interfaces for cyber-physical systems. *at - Automatisierungstechnik* 63, 10 (Oct. 2015), 833–843. https://doi.org/10.1515/auto-2015-0016

[23] Emese Papp, Christian Wölfel, and Jens Krzywinski. 2020. Acceptance and User Experience of Wearable Assistive Devices for Industrial Purposes. In *Proceedings of the Design Society: DESIGN Conference*, Vol. 1. Cambridge University Press, 1515–1520. https://doi.org/10.1017/dsd.2020.319

[24] Ana. C. Pereira and Fernando. Romero. 2017. A review of the meanings and the implications of the Industry 4.0 concept. *Procedia Manufacturing* 13 (2017), 1206–1214. https://doi.org/10.1016/j.promfg.2017.09.032

[25] Sandra Robla-Gomez, Victor M. Becerra, Josè Ramòn Llata, Esther Gonzalez-Sarabia, Carlos Torre-Ferrero, and Juan Perez-Oria. 2017. Working Together: A Review on Safe Human-Robot Collaboration in Industrial Environments. *IEEE Access* 5 (2017), 26754–26773. https://doi.org/10.1109/ACCESS.2017.2773127

[26] Arne Seeliger, Gerrit Merz, Christian Holz, and Stefan Feuerriegel. 2021. Exploring the Effect of Visual Cues on Eye Gaze During AR-Guided Picking and Assembly Tasks. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00041

[27] Arne Seeliger, Torbjörn Netland, and Stefan Feuerriegel. 2022. Augmented Reality for Machine Setups: Task Performance and Usability Evaluation in a Field Test. *Procedia CIRP* 107C (2022), 573–578.

[28] Jane Siegel and Malcolm Bauer. 1997. Field usability evaluation of a wearable system. In *International Symposium on Wearable Computers, Digest of Papers*. 18–22. https://doi.org/10.1109/iswc.1997.629914

[29] David Sousa Sousa Nunes, Pei Zhang, and Jorge Sa Silva. 2015. A Survey on human-in-The-loop applications towards an internet of all. *IEEE Communications Surveys and Tutorials* 17, 2 (2015), 944–965. https://doi.org/10.1109/COMST.2015.2398816

[30] Anna Syberfeldt, Oscar Danielsson, Magnus Holm, and Lihui Wang. 2015. Visual Assembling Guidance Using Augmented Reality. *Procedia Manufacturing* 1 (2015), 98–109. https://doi.org/10.1016/j.promfg.2015.09.068

[31] Philipp Url, Wolfgang Vorraber, and Johannes Gasser. 2019. Practical insights on augmented reality support for shop-floor tasks. *Procedia Manufacturing* 39, 2019 (2019), 4–12. https://doi.org/10.1016/j.promfg.2020.01.222

[32] Niels van Berkel, Jorge Goncalves, Simo Hosio, and Vassilis Kostakos. 2017. Gamification of Mobile Experience Sampling Improves Data Quality and Quantity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–21. https://doi.org/10.1145/3130972

[33] Susanne Vernim and Gunther Reinhart. 2016. Usage Frequency and User-Friendliness of Mobile Devices in Assembly. *Procedia CIRP* 57 (2016), 510–515. https://doi.org/10.1016/j.procir.2016.11.088

[34] W Wanbin and P W Tse. 2006. Remote machine monitoring through mobile phone, smartphone or pda. In *Engineering Asset Management*. Springer, London, UK, 309–315.

[35] Wenchao Wu, Yixian Zheng, Kaiyuan Chen, Xiangyu Wang, and Nan Cao. 2018. A Visual Analytics Approach for Equipment Condition Monitoring in Smart Factories of Process Industry. *IEEE Pacific Visualization Symposium* 2018-April (2018), 140–149. https://doi.org/10.1109/PacificVis.2018.00026

[36] Panpan Xu, Honghui Mei, Liu Ren, and Wei Chen. 2017. ViDX: Visual Diagnostics of Assembly Line Performance in Smart Factories. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 291–300. https://doi.org/10.1109/TVCG.2016.2598664

[37] Steffen Zenker and Sebastian Hobert. 2020. Design and implementation of a collaborative smartwatch application supporting employees in industrial workflows. In *27th European Conference on Information Systems - Information Systems for a Sharing Society, ECIS 2019*. Stockholm & Uppsala, Sweden, 0–16.

[38] Xianjun Sam Zheng, Patrik Matos da Silva, Cedric Foucault, Siddharth Dasari, Meng Yuan, and Stuart Goose. 2015. Wearable Solution for Industrial Maintenance. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 311–314.

[39] Fangfang Zhou, Xiaoru Lin, Chang Liu, Ying Zhao, Panpan Xu, Liu Ren, Tingmin Xue, and Lei Ren. 2019. A survey of visualization for smart manufacturing. *Journal of Visualization* 22, 2 (2019), 419–435. https://doi.org/10.1007/s12650-018-0530-2

[40] Fangfang Zhou, Xiaoru Lin, Xiaobo Luo, Ying Zhao, Yi Chen, Ning Chen, and Weihua Gui. 2018. Visually enhanced situation awareness for complex manufacturing facility monitoring in smart factories. *Journal of Visual Languages and Computing* 44 (2018), 58–69. https://doi.org/10.1016/j.jvlc.2017.11.004

Received February 2022; revised May 2022; accepted June 2022