ORIGINAL ARTICLE

Quantified factory worker: field study of a web application supporting work well-being and productivity

Päivi Heikkilä[1] · Anita Honka[1] · Eija Kaasinen[1] · Kaisa Väänänen[2]

Abstract

The work on the factory floor is gradually changing to resemble knowledge work due to highly automated manufacturing machines. In the increasingly automated work environment, the machine operator's task is to keep the production running and to solve possible problems quickly. This work is expected to become more autonomous, which raises the importance of supporting the workers' well-being. An important aspect of that is giving concrete feedback of success at work as well as feedback on physical and mental load. We implemented a smartphone optimized web application, Worker Feedback Dashboard that offers feedback to machine operators about their well-being at work and personally relevant production data as well as their connections to each other. The feedback is personal and based on objective, near real-time measurements. We present the results of a field study, in which ten machine operators used the application for 2–3 months. We studied the operators' user experience, usage activity, perceived benefits and concerns for the application with questionnaires, interviews and application log data. The operators found the feedback interesting and beneficial, and used the application actively. The perceived benefits indicate impacts on well-being as well as on work performance. Based on the results, we highlight three design implications for quantified worker applications: presenting meaningful overviews, providing guidance to act based on the feedback and refraining from too pervasive quantification not to narrow down the meaningful aspects in one's work.

Keywords Human-centered design · Field study · User experience · Factory workers · Quantified self · Operator 4.0

1 Introduction

The fourth industrial revolution, referred to Industry 4.0, is already on its way to factories, showing in increasing digitalization and automatization. Industry 4.0 is assumed to change the work on the factory floor towards knowledge work, requiring problem-solving skills and management of complex processes (Gorecky et al. 2014). The work on the factory floor may become more interesting, but also more complex, which may add to the mental load of workers and make it more difficult to stay aware of one's performance and development at work. Already today, a single operator may be responsible for an automated production line, which requires quick problem solving to avoid or minimize idle time in manufacturing.

To empower factory workers to receive encouraging feedback, to see their development at work, and to pay attention to their well-being along with work results, we designed and implemented a smartphone optimized, web-based Worker Feedback Dashboard application. The application provides factory workers with daily data-driven feedback on their well-being and work achievements based on measurements tracked with an activity wristband and retrieved from the production line of the factory. The application is based on the idea of quantified worker, which refers to the practice of self-tracking in the context of work. The feedback is intended for the personal use of the worker and it is not meant to be shared with the factory management.

✉ Päivi Heikkilä
  paivi.heikkila@vtt.fi

  Anita Honka
  anita.honka@vtt.fi

  Eija Kaasinen
  eija.kaasinen@vtt.fi

  Kaisa Väänänen
  kaisa.vaananen@tuni.fi

[1] VTT Technical Research Centre of Finland Ltd, P.O. Box 1300, 33101 Tampere, Finland

[2] Tampere University, P.O. Box 589, 33101 Tampere, Finland

In this paper, we present the results of a field study in which ten factory workers used the Worker Feedback Dashboard application for 2–3 months. The aim of this research was to understand the factory workers' user experience and usage activity in long-term use, as well as to study perceived benefits and possible concerns regarding the application.

The research questions of the study are:

- RQ1: How is Worker Feedback Dashboard used and experienced by factory workers in long-term use?
- RQ2: What kind of benefits and concerns do workers perceive from using the application?
- RQ3: What issues should be considered in the design of quantified worker applications?

In this paper, we first give an overview of the related research on health and well-being tracking at work and the vision of Operator 4.0. Second, we describe the Worker Feedback Dashboard application design as well as the field study method and participants. Then, we present the results of the field study and analyze the results to propose key design implications for quantified worker applications. Finally, we discuss wider implications of the quantification of workers.

2 Related work

The following two sections focus on the two main areas of related work relevant to our research: research on self-tracking, particularly in a work context, which has mainly focused on health perspective and the vision of the factory floor worker of the future—Operator 4.0.

2.1 Health tracking at work

Digital health and well-being applications have become increasingly common, both in leisure contexts and at work. The use of consumer wearable devices, particularly wristbands and smartwatches, has been increasing during the past five years (e.g. Rock Health et al. 2019), and recently also smart rings have gained interest (e.g. Oura, Moodmetric). Common health metrics tracked by these devices include physical activity, energy expenditure, sleep, and heart rate (Pardamean et al. 2020). More recently, also heart rate variability (HRV) and electrodermal activity (EDA) derived stress-related metrics have been included in many consumer devices.

Utilization of personal health tracking technologies has become an integrated part of a growing number of workplace health and wellness programs (Moore and Piwek 2016). The practice of personal health and well-being tracking is based on the Quantified Self trend (Wolf 2010), which

refers to self-monitoring of biological, physical, behavioral, or environmental data (Swan 2013). The goal of the Quantified Self practice is to gain meaningful insights from the self-tracking data, which enables positive changes in behavior (Choe et al. 2014). Also other terms, such as life-logging [see Gurrin et al. 2014; Selke 2016], personal analytics (MIT Technology and Review 2020) and personal informatics (Li et al. 2010) are used about the same phenomenon of enhancing self-knowledge by recording personal data. We use the concept of the quantified worker to refer to the practice of personal self-tracking in a work context, which may be enabled or proposed by the employer. Thus, the term quantified worker solution refers to a digital solution including self-tracking data, such as a data-driven feedback application.

Promoting health and well-being tracking at work may come with many benefits, such as reducing the physical and cognitive burden of the workers (Lavallière et al. 2016), early detection of health problems (Li et al. 2017) and fostering healthy behavior (Asimakopoulos et al. 2017). The benefits include increasing awareness of one's daily activity and personal accountability towards health goals (Chung et al. 2017). Wearable trackers have a potential to provide a new kind of worker feedback, which is personal, immediate and objective (Piwek et al. 2016). For the factory floor workers of the future, wearable trackers are expected to have the potential in supporting the workers' occupational health, safety and productivity (Romero et al. 2018).

Despite the potential benefits, health and well-being tracking involve a variety of ethical issues, such as concerns related to privacy, data security or true voluntariness of self-tracking (Heikkilä et al. 2018b; Lupton 2016; Moore and Piwek 2016). Furthermore, receiving personally meaningful insights through self-tracking may be challenging. Self-tracking at work has been found to focus on easily measurable metrics, and not on supporting personal health goals or well-being in a more holistic way (Chung et al. 2017). This may decrease long-term motivation for self-tracking. In one qualitative study (Masson et al. 2016), activity trackers were given to 13 users at the same workplace, and they all stopped using them within three months. One reason for this was the limitations of the current technology, for example frustration caused by mechanistic feedback of the device due to simplistic models of data analysis. The studies among active self-trackers have identified also other pitfalls or barriers of self-tracking, such as tracking fatigue due to tracking too many things (Choe et al. 2014), not receiving meaningful insight due to lack of contextual data (Choe et al. 2014; Li et al. 2010) as well as insufficient motivation, lack of time or forgetting to self-track (Li et al. (2010).

In contrast to the identified barriers, providing visualizations and contextual information have been found to make the data more meaningful to self-trackers, increasing a sense of accomplishment (Pantzar and Ruckenstein 2015). When

activity-tracking practices were studied with video methods (Gouveia et al. 2018), the study revealed the relevance of providing glanceable information, which enables immediate learning of the information, as well as facilitating of micro-plans, such as reaching 1000 steps in the next hour, instead of deep retrospective analysis of self-tracking data. Supporting these practices with design could be applicable to the work context as well, as long as self-tracking does not provoke too frequent checking of data not to distract the user from the work tasks.

2.2 Operator 4.0

Industry 4.0 will radically change many work roles in the industry. For the industrial workers, the revolution is expected to provide opportunities by the qualitative enrichment of their work: a more interesting working environment, greater autonomy and opportunities for self-development (Gorecky et al. 2014). The change in the factory floor work has been characterized as Operator 4.0 (Romero et al. 2016). Operator 4.0 vision refers to smart and skilled factory operators of the future, who are assisted by automated systems, allowing the operators to utilize and develop their creative, innovative and improvisational skills, without compromising production objectives (Romero et al. 2016).

While work well-being at the factory floor has earlier been focused on physical fatigue, Operator 4.0 raises well-being issues related to mental fatigue. Cognitive workload can be relieved by assessing the operator's well-being at work by using wearable technology (Romero et al. 2018). In the increasingly automated work environments, factory floor workers may miss concrete feedback of their work achievements and competence development. As awareness of one's work performance has an impact on job motivation (Hackman and Oldham 1975), it is important to support receiving meaningful feedback on one's work. Meaningful feedback that supports the autonomy of workers by allowing them to decide how to utilize it has the potential to support their intrinsic motivation to engage in work activities instead of having to engage in them [see e.g. Gagné and Deci 2005]. In a wider scope, workers' intrinsic motivation can be supported with autonomy-supportive work environments and with interesting and challenging jobs that allow choice (Gagné and Deci 2005).

Even though tracking of health and well-being at work has become common, only a few empirical studies focus on the perspective of the workers (Chung et al. 2017), and the long-term impacts of self-tracking at work have not yet been systematically studied (Moore and Piwek 2016). Combining tracking of well-being related metrics and work performance-related metrics is a new approach, which evokes new research questions and ethical issues. To the authors'

knowledge, a similar solution has not been piloted in factories or other work contexts.

3 Worker Feedback Dashboard

In this section, we first explain the design rationale of the Worker Feedback Dashboard web application [for more details of the design process and design decisions, see (Heikkilä et al. 2018a)]. Then, the structure and content of the application are presented with illustrations of the application views.

4 Design rationale

We designed the Worker Feedback Dashboard web application to offer factory floor workers a new kind of data-driven near real-time feedback of metrics related to their well-being and work achievements. The application provides a possibility to discover connections between well-being data and personally relevant production data and highlights positive progress in work performance. It shows both work shift-specific metrics and trends over a longer usage period. The web application is optimized for mobile devices, and it is responsive to different screen sizes. To preserve the workers' privacy, only the worker sees one's personal data and metrics in the application.

The concept of the Worker Feedback Dashboard was developed based on the insights from interviews with factory workers (Heikkilä et al. 2018a). The design process followed a human-centered design approach (International Standardization Organization (ISO) 2010), the principles of a participatory design process (Kuhn and Muller 1993; Schuler Namioka 1993) and the Worker-centric design and evaluation framework for Operator 4.0 solutions (Kaasinen et al. 2018). The ideology of Quantified Self, receiving meaningful insights from self-tracking data (Choe et al. 2014), and the vision of Operator 4.0, smart and skilled operators of the future (Romero et al. 2016), were applied as a starting point for the design.

The application was designed to respond to four user experience goals [see Kaasinen et al. 2015; Roto et al. 2017] defined during the design process: (1) being empowered and encouraged, (2) getting personal feedback, (3) getting meaningful insight and (4) being undisturbed (Heikkilä et al. 2018a). The application aims at encouraging and empowering workers by highlighting positive aspects and achievements of the work shift, and at providing meaningful insight through enabling personal reflection of the data. It also aims at providing personal feedback, which offers an opportunity to have an impact on the factors tracked and the possibility to see one's development in work tasks. Personal feedback

means also that the application is intended for the personal use of the worker, not to be shared with the management or other employees. To avoid disturbing the users in their work tasks, real-time notifications were excluded, and the main view was designed to present the information so that it can be checked quickly, for example during a break at work.

4.1 Application structure and content

The feedback provided by the application is presented visually with graphs and charts. The main view of the user interface presents the data of one work shift, including well-being metrics, production metrics and a time-series graph showing selected well-being metrics together with the main production outcome. In addition, the longer-term evolvement of any of the metrics can be seen in a trend view. The trend view is separated from the application main view with a tab, to devote the main view for the daily data. Well-being metrics from an activity wristband and production metrics from the production line are retrieved every 20 min. All the data is stored in an external server and can be accessed only by the user.

The well-being metrics (Fig. 1), tracked with an activity tracker (Fitbit Charge 2 in the current implementation), are shown on the topmost part of the application's main view. The selected metrics include the quantity of steps during the work shift as well as the resting heart rate for the day. If the users use the tracker while sleeping, the view also shows the amount of restful sleep. These three metrics were included, as they give indications of the user's

activity, stress and recovery in an easily understandable format. Furthermore, seeing the number of steps at work and the restful sleep duration may encourage the user to make behavior changes for example to sleep more. In addition to these three metrics, the users are asked to evaluate their perceived concentration level at work for each work shift with a five-scale rating. The evaluation is asked to trigger users to engage with the application and to facilitate reflecting on the workday.

The production metrics, tracked from the production line, are shown in the middle part of the application main view. The metrics were co-designed with factory workers who are operating automated, multipurpose sheet metal processing lines. The production metrics are intended to support the work objectives of machine operators whose role is to keep the line running by monitoring the process and solving disturbances.

The selected production metrics (Fig. 2) include the utilization rate of the machine, the longest continuous machine-running period during the work shift and the average recovery time after failures (the time taken to resolve an error situation). The metrics are defined to highlight positive progress, for instance showing machine running time instead of idle time and highlighting the time for solving failures instead of the number of them. We consider this important, as the factory floor workers' interviews, conducted earlier during the research project, revealed that workers' work performance is typically reported through negative measures (Gorecky et al. 2014). In addition to the positive definition of metrics, positive

Fig. 1 The topmost part of the application main view presenting well-being metrics
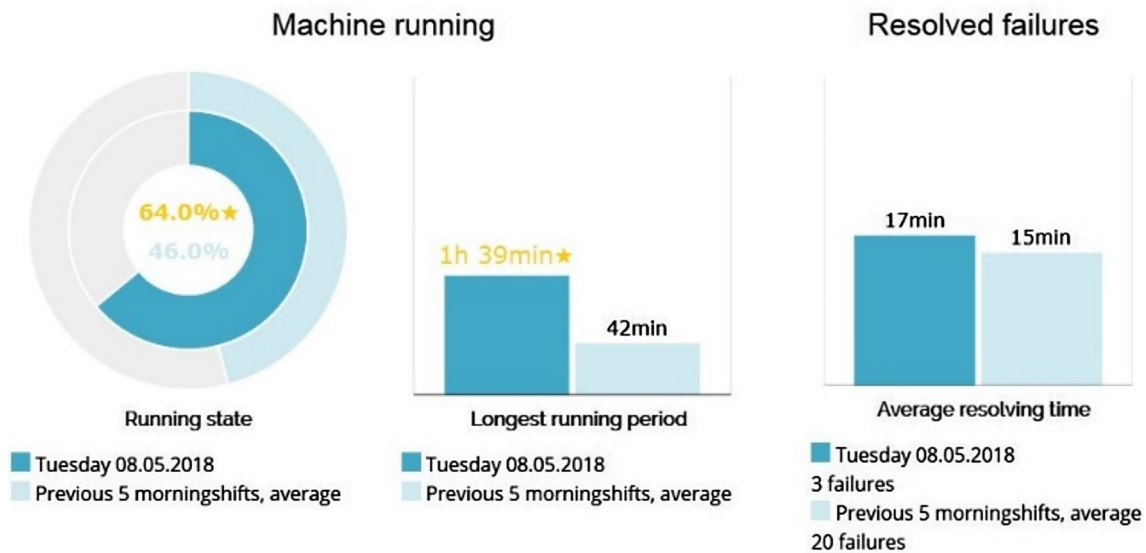
Fig. 2 The middle part of the application main view presenting production metrics during a work shift

progress of each metric, compared to the previous five work shifts, is notified with a star icon beside the value of the metric.

Below the production metrics in the application main view, a time series graph provides detailed information of the current or the latest work shift. It shows two selected well-being metrics, the user's heart rate and the quantity of steps, together with the main production outcome (Fig. 3), the utilization rate of the machine. The selected metrics are such that they vary during the workday and thus, help the user to recall what happened during the work shift and what kind of reactions different events evoked. This may facilitate the user to reflect on, for example, how one reacted in the error situations of the machine or whether one experienced stressful or recovering moments during the work shift. This, in turn, may give hints of beneficial ways to modify one's behavior at work.

A trend view (Fig. 4) can be accessed from the main view via a tab. The view enables discovering connections between metrics or seeing evolution of one metric over a longer usage period. The user can select three of the metrics to be presented simultaneously during a preferred timescale. The different types of work shifts (e.g. morning shifts and night shifts) are presented in the background with different colors. The view can be used, for example, to see whether the user has learned to resolve errors of the machine more quickly over time or whether one's self-reported concentration level seems to be connected to the amount of restful sleep.

5 Study design

We conducted a field study to understand the factory workers' user experience, usage activity, perceived benefits and possible concerns of the Worker Feedback Dashboard solution. The study is based on Research through Design approach (Zimmerman et al. 2007), in which the designed artifacts form a significant contribution of the research.

5.1 Participants

Ten participants from three metal industry factories participated in the study. All of them were machine operators of modern, highly automated, multipurpose manufacturing machines. The participants can be seen as early representatives of the Operator 4.0 vision (Romero et al. 2016), as their work requires independent problem solving and holistic understanding of the production process. All operators who used the machine in these factories were invited to participate in the study and they all wanted to take part. The majority of the participants were males (7/10), with a mean age of 32 years, ranging from 22 to 50 years. Three of the participants operated the multipurpose manufacturing machine connected to the Worker Feedback Dashboard full-time, while others also used other machines or worked as machine programmers as a part of their work. Most of the participants had worked in their current role for less
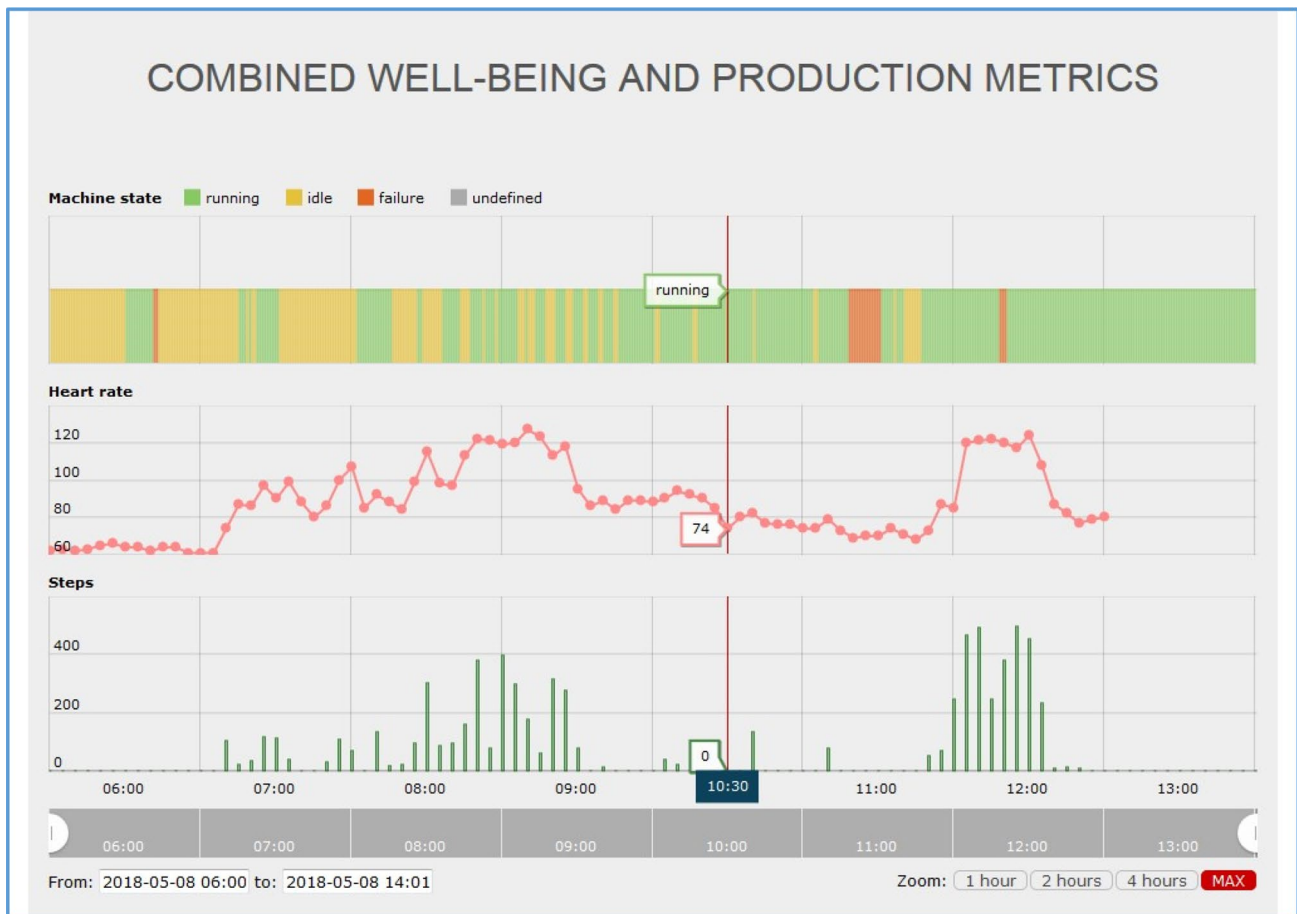
COMBINED WELL-BEING AND PRODUCTION METRICS

Fig. 3 The lowest part of the application main view presenting a time-series graph of selected well-being metrics and the main production metric

than two years. Half of the participants had used a well-being tracker before.

5.2 Procedure and data gathering methods

In each of the three factories involved, all machine operators of the multipurpose machine were first invited to a session where we introduced the Worker Feedback Dashboard and the study protocols to them. The participants were explained the aim of the research, the overall purpose of the application, practical steps of the participation, data management practices of the study and one's rights to refuse from participation or withdraw from the study at any stage without negative consequences. We introduced the participants the main functionalities of the application but avoided guiding them too much. Both well-being and production metrics of the application were briefly discussed with the attendees. Related to well-being metrics, special attention was paid to the metric of resting heart rate to inform the participants that it may indicate stress but may vary also because of other reasons. Concerning production metrics, the objectives were briefly discussed to ensure that the workers understand that

the metrics give feedback of each workday, but do not indicate that reaching the maximum number of a metric (e.g. 100% utilization rate of the machine) would be possible or expected from the workers.

After the introduction, the attendees of the session could choose if they wanted to participate in the study. The management of the factories was not present, to ensure a more unbiased decision of participation. The workers who wanted to participate signed the informed consent forms and were encouraged to ask any further questions related to the participation. They could also choose which well-being-related data (data on sleep, heart rate and steps) they allow to be used by the feedback application. We helped the participants to take the application and the activity wristband into use and instructed them to use the application as part of their work and everyday life for the study period of 2–3 months. We recommended that they would use the activity wristband at least while working (Fig. 5) and that they would check the application regularly during the first two weeks of the study, preferably during or after each work shift. The participants were informed that they could keep the Fitbit device that was provided for them for the study if they used the application
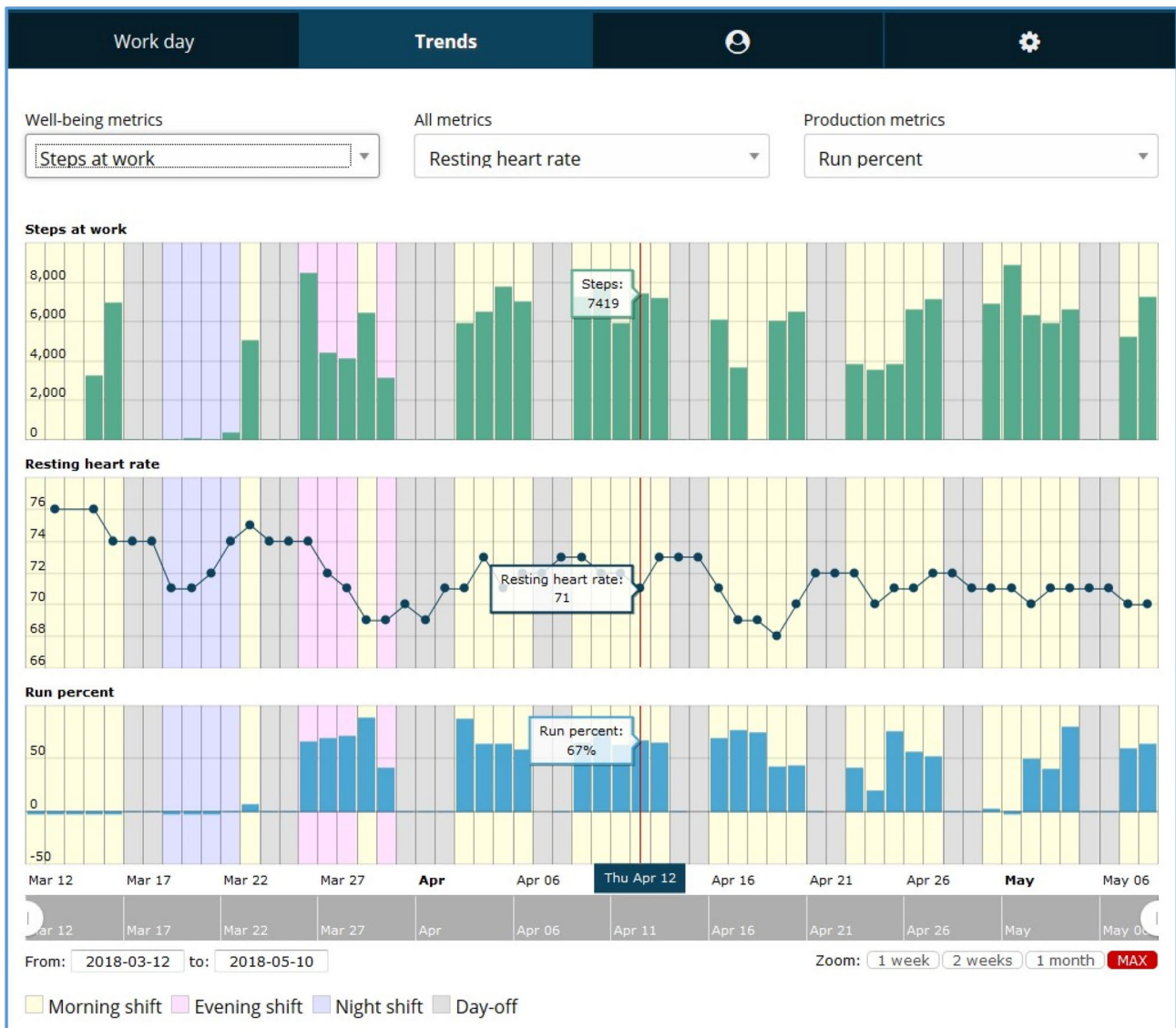
Fig. 4 Trend view of the application presenting longer-term trends according to the user's selections

for at least for two weeks. During the study, the participants had an opportunity to ask help from the researchers by phone or by email if they had any technical or other problems with using the Worker Feedback Dashboard application or the Fitbit device.

After using the application for about two weeks, the participants received an online questionnaire to study their ways of using the application and first impressions of it. At the end of the 2–3 month usage period they received a final questionnaire of their overall user experience. Nine participants filled in the first online questionnaire and eight participants filled in the final questionnaire. Finally, eight participants attended a 1 h individual theme interview that focused on the perceived benefits of the usage of the solution and potential concerns of the participants. In addition

to the questionnaire and interview data, application log data were collected regarding the usage activity. The focus of this study was user experience and user acceptance of the Worker Feedback Dashboard. Thus collecting objective data from the wristband or the production system was out of the scope of this study.

When designing the questionnaires and the structure of the interview, Worker-centric design and evaluation framework for Operator 4.0 solutions (Kaasinen et al. 2018) was utilized. The framework covers both immediate implications and wider impacts. Immediate implications include five aspects: user acceptance, user experience, usability, safety and ethics. Each of the aspects were studied in the questionnaires with one or more statements or questions. User acceptance was studied focusing on interest towards

Fig. 5 A factory worker operating a highly automated manufacturing machine while wearing an activity wristband for tracking well-being metrics

the provided data and perceived usefulness of the application (the content of the app is interesting; the app seems useful). User experience was studied through the overall feeling of the usage and perceived pleasantness of it (indicate your overall feeling of using the solution by choosing the best matching expression of smiley faces; using the app feels pleasant). Usability was evaluated through three statements (navigation in the app feels effortless; the structure of the app is clear; it is easy to understand the information provided). The aspect of safety was assessed through two statements (using the app hampers working; using the app at work has caused safety hazards), and finally, ethics was studied through one generic statement (using the solution at work feels questionable).

The framework (Kaasinen et al. 2018) studies foreseen impacts regarding work well-being and productivity. In this study, the impacts were studied through perceived benefits reported by the users in the interviews. In addition, the questionnaire included one general impact-related statement (The application has had an impact on my daily activities).

5.3 Data analysis

The questionnaire data of the study were analyzed mainly quantitatively. In this paper, we report the main results of eight respondents to the final questionnaire. In addition, the reported results include two questions related to usability from the first questionnaire (clarity of the structure of the application and effortlessness of the navigation), as those questions were not repeated in the final questionnaire. User experience and user acceptance of the worker Feedback Dashboard did not differ much between the first questionnaire (after 2 weeks of use) and the final questionnaire. For example, all operators considered the same content of the

application as interesting based on the use of two weeks as well as in the end of the usage period. Because of this and due to our focus on the long-term impacts, we report only the results of the final questionnaire. Questions related to safety and ethics were reverse scored to present the results of immediate implications in a way that the most positive option of the Likert scale always refers to the most positive rating of the implication (e.g. concerning ethics, the reversed statement refers to an ethically sound solution: using of the solution at work does not feel questionable).

The interviews were analyzed qualitatively by focusing on perceived benefits and concerns related to the usage of the application. The results include illustrative citations from users describing the perceived benefits and concerns derived from the eight user interviews (U1–U8, U referring to a user). Log data were analyzed to extract metrics for describing usage activity based on identified usage sessions.

6 Results

In the following, we respond to the research questions of the study:

- RQ1: How is Worker Feedback Dashboard used and experienced by factory workers in long-term use?
- RQ2: What kind of benefits and concerns do workers perceive from using the application?
- RQ3: What issues should be considered in the design of quantified worker applications?

We first present the results of usage activity, user evaluation of the application, perceived benefits as well as users' concerns and ethical considerations related to the application. Then, we reflect the results on user experience goals for the application and propose three key design considerations for the design of data-driven worker feedback applications.

6.1 Usage activity

According to the log data, five of the ten participants used the Worker Feedback Dashboard on at least on 50% (range 15.5–74.7%) of the days included in the study period. The number of workdays which involved operating the multi-purpose machine varied from four to forty during the study period for different participants (most participants had also other responsibilities and workdays when they did not operate the machine). When considering the number of the application usage days in relation to the number of work shifts that involved machine operation, for nine of the ten participants the number of usage days was at least 0.5-fold compared to the number of these work shifts. Thus, it seems that most of the participants used the application on at least

every other workday that included machine operation. Half of the participants used the application more frequently than the number of workdays involving machine operation. For these participants, the ratio between the number of usage days and the number of work shifts involving machine operation varied from 1.7 to 3.3-fold, which means that the application was used also during other types of work shifts or on days off. According to the self-reported usage activity, participants typically checked the application once or twice a workday, during breaks or after the work shift.

6.2 User evaluation of the application

The participants' overall experience of the application was studied through questionnaire statements assessing the five aspects of the Evaluation Framework's (Kaasinen et al. 2018) immediate implications. According to the responses, the immediate implications related to the usage of the application were evaluated as positive (Fig. 6). The content of the application was considered as interesting and the application seemed useful to the participants (user acceptance). The user experience and usability were identified as positive. Usage of the application had not caused danger or hampered working (safety), and the usage of the application was not felt to be ethically questionable (ethics).

In addition to the overall evaluation of the application, we studied how interesting the participants found the content provided (Fig. 7). The most interest was shown towards the amount of restful sleep, which was found very interesting by most participants and interesting by all. The users were also interested in the number of steps taken during the work shift, resting heart rate, the utilization rate of the machine and the graph of combined well-being and production measures (see Fig. 3). However, the information on the longest continuous

run of the machine and the recovery time from failures were not perceived as interesting by everyone. This was confirmed in the user interviews: some found these metrics interesting and encouraging, but they were not considered relevant in all factories. Similarly, not everyone found the trends view interesting. The interviews revealed that some of the users had not checked the trends at all, possibly because the information was not located in the main view and thus could be easily overlooked. Self-reporting of the concentration at work was considered as the least interesting content. Not all participants reported the value, and from those who did, not all had checked the trend information, which could have made this value more insightful.

6.3 Perceived benefits

The foreseen impacts of using the application were studied by interviewing the participants about the perceived benefits of the application. In the questionnaire, half of the respondents evaluated that the usage of the application had an impact on their daily activities. The interviews revealed that the perceived benefits were related to both receiving personally relevant production data and positive feedback as well as receiving data related to well-being and seeing its connection to work performance. In the following, we describe first the benefits supporting productivity and then the well-being-related benefits. However, in most cases, these are intertwined. For example, motivating feedback may have an impact on well-being and work performance, which may lead to enhanced productivity.

Personal production data helped the participants recognize and follow their accomplishments at work. The machine utilization rate was interesting for almost all workers, as it was regarded as a concrete indicator of work results: "When
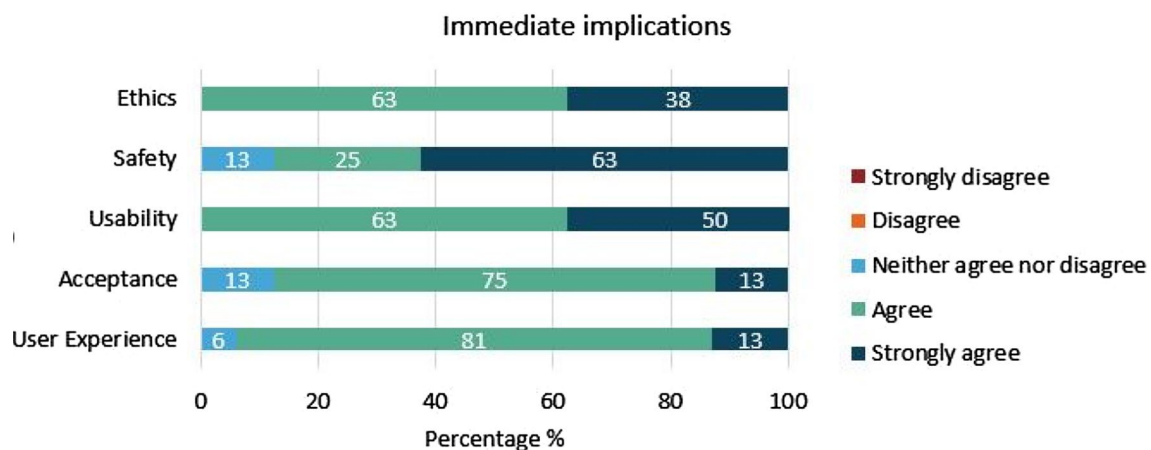
Fig. 6 User evaluation of the immediate implications of the application, presented as a distribution of the users' responses to the statements related to the five aspects of the evaluation framework (Sect. 4.2). The "strongly agree" option refers always to the most positive rating of the dimension (see Sect. 4.3)

Fig. 7 Participants' interest towards the content types of the Worker Feedback Dashboard application

Interest towards the content of the application
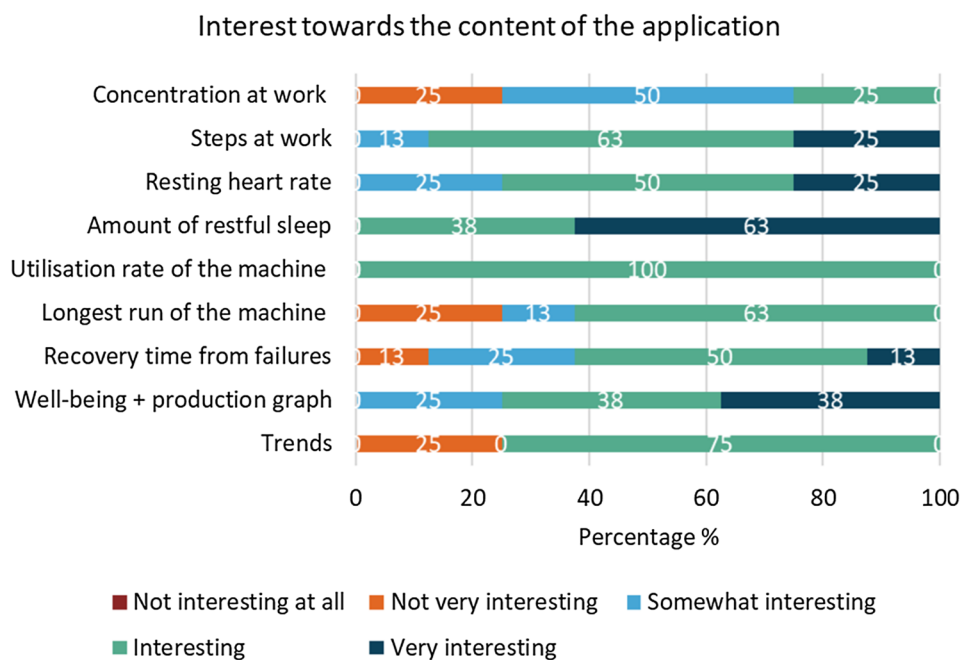
the machine utilization rate is high, you get a feeling that you have accomplished something" (U6). The value was especially interesting for those workers who had this metric explicitly prioritized in their work: "I'm most interested in the machine utilization rate that allows me to see what I have accomplished at work. Our bonus depends on that. It is my first priority at work and it is easy to have an impact on, through the order of (production) tasks or by preparing for adding the next sheets to the machine." (U2).

Personal production feedback not only provided a means to improve one's performance but also motivation for better performance, as it concretized the impact of one's efforts: "I perhaps try a bit more at work as I like to compete (laughs). I already knew the ways to do things faster, but haven't put effort into that. As you can't see it anywhere." (U1) Quantification also enabled the setting of measurable targets for oneself: "I wanted to get five hours (machine running time) each day" (U8). In addition to daily feedback, also summarized feedback over a longer period, such as weekly summaries of the data, was proposed to be included in the design to give an overview of the feedback.

Users valued positive feedback from the application. Participants liked the star indications showing positive progress of a metric, and two participants commented that they could have been even more visible in the user interface. Two participants commented that they liked the metric of the constant run of the machine, as it provided new uplifting information: "It was fun to see the longest run (of the machine) as you would not think about it otherwise" (U6).

The new kind of feedback facilitated self-reflection and provided ideas for developing the ways of working. In particular, the detailed graph of the workday with steps, heart rate and the state of the machine helped to recall what had happened during the work shift. It was used for self-reflection and finding evidence for one's own feelings: "When I felt that the workday was hard, I checked from the data whether it was reflected there (e.g. in the step count)" (U3). Even though not all participants reported their perceived ability to concentrate on work tasks to the application, those who did, considered the reporting activity as a good daily trigger to access the application and a basis for self-reflection. For one user, a high quantity of steps during the workday indicated that the work could be organized in a more optimal way: "I was surprised about the quantity of steps during the workday. This means that there would be an opportunity for development as the materials needed haven't been easily available, and you have had to search for them." (U3) Another user noticed that the days when he was able to concentrate on one work task at a time and used only one machine led to better results than the days involving several tasks or machines.

Data on well-being, especially related to sleep, encouraged the participants to pay attention to the amount of sleep and recovery between workdays. The data on sleep was found interesting both by the participants working in three work shifts as well as the participants having only daytime shifts. One participant noticed a connection between sleep and work performance: "If you sleep badly the day is a mess. When you slept longer the day was better." (U2) However, it was typical to discover that the amount of sleep was quite low, but the users had not necessarily changed their behavior to sleep more. Thus, raising awareness of the potential

means to change one's behavior could make the application more beneficial. In addition to monitoring sleep, the usage of the fitness tracker encouraged more regular exercising, which is likely to have a positive impact on work well-being as well.

6.4 Concerns and ethical considerations

Ethical considerations were studied by discussing the potential concerns of the participants during the interviews. Even though the participants had not evaluated the usage of the application as questionable in the final questionnaire (Fig. 6), we wanted to further elaborate on the subject with them.

Based on the interviews, the main concerns of the participants did not relate to privacy or data security, but wider themes covering the relevance and the purpose of the application. As the work of the participants was versatile and the work tasks variable due to the nature of different orders and parts to be produced, the production metrics and progress indicators were not always realistic or meaningful to workers. Even though the production parameters of the application were co-designed with factory workers of one factory and all the workers had experience on operating the exactly same machine, the prioritization of metrics and the nature of the work varied between the factories. For example, long-running periods may be impossible to achieve when the work requires short operating times due to small production batches: "In the day time you need to do the bits and pieces. You would need to set the machine to run slower to get longer runs (laughs). The easier work and the longer runs are done in the evening shifts." (U4) This means that encouraging the users to achieve as long continuous operating time as possible can be irrelevant or unrewarding when the most demanding work tasks require shorter operating times.

Even though the users mainly considered the quantification relevant, they wanted to highlight that this kind of application cannot give a complete picture of the work performance, as the work also includes tasks and goals that cannot be measured by the operation of the machine. The work includes more qualitative aspects as well as versatile work tasks while operating the machine: "You can't see what you do when the machine is running. Like when you fill in or offload the sheet supply. Or when you use the truck, it doesn't accumulate your steps. So you can't directly see your activity from the step count." (U3) In addition, one user emphasized that the running percentage of the machine is not relevant if the manufactured parts are not of high quality: "Even though the machine might be working seven hours a shift, it doesn't tell the whole truth. For example, if the parts are good only for a trash pick-up." (U8).

The participants did not personally find the application questionable, but they assumed that the application would not be accepted by all workers: "Some would like to use it but some would say they are definitely not wearing anything. They would see it as a way to be controlled." (U1) The users recognized that the application could be misused and it could be used unethically to compare the work performance of workers if the data was available to the management. Despite this, the users did not raise privacy as a significant ethical issue. However, possible misuse of personal data in similar services was discussed in the interviews and the negative consequences were brought up: "The data is like in a vault on the Internet. You never get it away from there." (U5).

7 Design considerations

7.1 Reflection on user experience goals

The results reveal the applicability of the four user experience goals set for the application when designing it. The goals were (1) Being empowered and encouraged, (2) Getting personal feedback, (3) Getting meaningful insight, and (4) Being undisturbed. The users considered all these aspects relevant. They were reflected in the benefits that the participants had perceived, and the users found them worth pursuing when we explicitly asked this in the final interviews.

First, positive production metrics and highlighting of work accomplishments motivated workers and created an uplifting feeling of working. Thus, the use of the application could be seen as encouraging and empowering for the workers. The participants wished the positive indications to be even more visible in the user interface. Second, the feedback provided was found to be personal. The feedback motivated the participants to improve their performance by concretizing the impact of one's efforts and creating a possibility to compete with oneself. In addition, it encouraged setting personal goals for the workday even though it was not an explicit feature of the application. Third, the users considered the provided feedback mainly meaningful, but this could be further enhanced by letting the users customize the metrics shown, by promoting the possibility to see longer-term trends of the data and by providing recommendations and conclusions based on the data. We highlight these aspects in the design implications proposed for providing meaningful overviews and guiding the user to act based on the feedback. Finally, the users did not feel that the application disturbed their work. They mainly used it during breaks or after the workday and did not find the usage distracting them from the work tasks. When we discussed a possibility to add real-time notifications to the application, most of the participants preferred the summarized feedback to any real-time feedback.

7.2 Design implications

The results of our field study give further insight into developing quantified worker solutions. We propose three key design implications for the design of data-driven worker feedback applications. We summarize these from the worker point of view as (1) give me meaningful overviews, (2) guide me to act based on the feedback, and (3) do not underestimate the unquantified.

7.2.1 Give me meaningful overviews

To support users in finding relevant information at a glance, giving personally meaningful overviews is important. To increase the meaningfulness of the data, the user should be able to choose the metrics to be followed also in the main view and be able to check both work shift specific feedback as well as longer-term feedback easily.

In our trial, the trend view was used only rarely by a few participants. As it is not a relevant feature at the very beginning of the usage when no data has been accumulated yet, a notification of it after two or more weeks of usage could remind the users of the feature, and thereby, enable the discovery of meaningful connections in the data. In similar trials, notifying the users to try out features that become topical after more data has been cumulated might increase the insights received and thus, provide additional benefit to the users. A more advanced, useful feature would be to analyze patterns in the data in the background, and notify the users when new patterns are detected.

Another means to increase meaningfulness is to show the user relevant metrics that are either important in one's work or personally interesting for the user. In our study, we found out that the nature of the work and the priorities were different in different factories, even though the users of the application operated the same machine. Thus, the same metrics were not relevant for all. To provide meaningful data for the user, the application should be customizable for each workplace, and in addition, the user should be able to choose the metrics to be included.

7.2.2 Guide me to act based on the feedback

Meaningful overviews may provide interesting information for the user, but the feedback becomes more valuable if it helps the user in finding ways to put the knowledge into practice. In our study, the content of the application was considered interesting and the participants perceived benefits of the usage, but half of them stated that the application did not have an impact on their daily activities. Instead of providing information only, recommendations or conclusions based on the data might increase the application's potential to facilitate positive behavior changes and thus, its

impact on everyday life and work. For example, providing insights on one's progress at work tasks during a longer-term period, allowing the worker to set personal goals or providing behavior change tips could facilitate finding new ways of working that would lead to better results. Besides work goals, the same concerns well-being. Currently, the application helps the user to see, for example, the benefits of having enough restful sleep, but it does not give advice on how to improve sleep or avoid sleeping problems. In the future, artificial intelligence could be utilized to provide personalized behavior tips that would motivate the user to act based on the received feedback.

7.2.3 Do not underestimate the unquantified

The third key design implication is based on the concern that a data-driven worker feedback application may create a feeling for the worker that the scope and diversity of one's work is narrowed down to simple metrics. To avoid that, quantification should not claim to capture everything the user does during the workday. Although the increasing utilization of artificial intelligence and emerging tracking technologies may create a temptation of quantifying more and more aspects of the user's workday and life, leading to an illusion of a comprehensive view of the worker, striving for total pervasiveness could have several negative consequences. First, it would neglect some aspects of the work, as not everything can be quantified. Second, it might make the workers overvalue the metrics that are quantified at the cost of aspects that cannot be measured, such as helping co-workers. Third, it could make the work feel mechanistic, for example focusing on quantity over quality. Besides that, it could make workers more reluctant towards applications based on quantification, because pervasive quantification might convey an image of controlling the worker and create suspicion of the purpose of quantification. Hence, it is important to provide quantified feedback to be used as a support for self-reflection and the worker's intrinsic motivation to consider well-being or to improve one's ways of working, not for measuring the worker's performance or pursuing to turn all aspects of the work into numbers.

8 Discussion

This field study explored the overall experience, perceived benefits and participants' concerns related to the Worker Feedback Dashboard solution. According to the results of the long-term field study, the immediate implications of the application usage were considered positive and the users perceived benefits related to personal well-being and awareness of work achievements. For example, the application helped the users recognize their accomplishments at work,

which increased motivation to make efforts to improve one's work performance. However, the perceived benefits included also remarks for improvements in designing and structuring the work practices at the workplace, which implies that the application could have a wider impact in the work community.

The derived design implications for the design of quantified worker solutions include giving meaningful overviews of the data, providing guidance for acting based on the data and defining the scope of and purpose of quantification carefully, not to give an impression of underestimating the unquantified aspects of one's work and narrowing down the versatility and diversity of it. These implications facilitate both discovering personally meaningful information as well as making positive behavior choices based on the feedback, still avoiding striving for the total pervasiveness of quantification. The identified design implications may help in overcoming some barriers to self-tracking found in earlier studies, such as tracking fatigue (Choe et al. 2014) or insufficient motivation (Li et al. 2010). Providing meaningful overviews and guidance to act support also learning through glances and facilitating of micro-plans (Gouveia et al. (2018), by enabling quick checking of the most interesting data and support for putting the knowledge into practice.

The participants of this study were not personally concerned about ethical issues, but they brought up that the relevance and purpose of the application should be well designed. Both the factory management and workers should be involved in defining the production objectives to ensure their relevancy in terms of production as well as workers' targets and tasks. In addition, all stakeholders should be aware of the purpose of the application, to avoid incorrect impressions of it. The workers should be aware of realistic expectations towards them, for example, that they are not expected to reach 100% utilization rate of the machine even though it is the theoretical maximum of the metric. If all information is available only for the workers themselves, the employer and the management should be aware of this. Potential wider impacts of the application use could be facilitated by providing selected anonymous information of bigger samples as summaries to the management or the employer. However, this might impair the workers' trust towards the application and blur the idea of it as providing feedback for the worker only. If data-driven feedback is used for wider purposes, the users should be aware of this and a neutral service provider could be used to convey the summaries to the employer. This role could be given for example to an occupational health agent, which could also help workers in interpreting the biometric well-being data if needed.

Although the study participants did not report any negative consequences of receiving data-driven feedback regarding their well-being and work performance, for some users unwanted effects might occur. If the feedback negatively conflicts with one's self-image (e.g. work performance metrics are not as good as one would have expected), this might decrease the feeling of competence and self-esteem (Stiglbauer et al. 2019). On a positive note, increased self-awareness of one's improvement needs may motivate individuals to take actions towards a healthier lifestyle or to seek opportunities for competence development (Stiglbauer et al. 2019).

The links between physiological data and mental states are not always straightforward. For example, heart rate may be an indicator of physical or mental load and it is related to attention as well (Vanderhaegen et al. 2020). Even though the interpretation of physiological data would not always be clear, it can provide interesting data for self-reflection. Still, the accuracy of the data is important to prevent misleading individuals. Consumer wearables generally tend to measure physical activity and heart rate with sufficient accuracies (Fuller et al. 2020), but for sleep metrics there is room for improvement (Zambotti et al. 2019). Furthermore, the impact of data-driven feedback on users' psychological state should be considered. Previous experiments show that merely providing negative feedback on one's sleep quality deteriorates mood and increases the feelings of fatigue, irrespective of the actual sleep quality (Gavriloff et al. 2018). A similar phenomenon might apply also to perceived stress, where personal measurements indicating high-stress levels might increase perceived stress. Hence, highlighting positive feedback and for example considering what would be the appropriate time of day to reveal such sensitive metrics to the users is important.

If the work performance is evaluated through quantified feedback, the meaning of non-quantified tasks may be underestimated, or the goals of the work may appear narrower than before. The same has been remarked in workplace wellness programs, which often aim at holistic health support but in practice tend to focus on or incentivize the number of steps only, as it can be easily tracked (Chung et al. 2017). Even though everything cannot be quantified—and is definitely not something to be striven for—the socio-technical gap between the goals to be supported and technical means to support them (Ackerman 2000) will potentially be bridged in the near future due to the possibilities of advanced artificial intelligence-driven data analysis. In the design of data-driven worker feedback solutions, the balance between quantifying relevant metrics in a holistic way, still avoiding total pervasiveness is a central design question.

Even though the application usage did not raise major ethical concerns in this study, the participants remarked that such solutions would not be accepted by all factory workers. Voluntary participation is a key principle that should be ensured when introducing similar solutions to workplaces. The workers should not feel coerced to participate and should be able to end participation at any time without

negative consequences. In addition, transparency in collecting the data may alleviate concerns related to potential misuse of the data. In practice, this can be implemented for example by giving the users a possibility to select which data (e.g. steps, heart rate, sleep data, production data etc.) they allow to be used by the application.

According to this field study, the Worker Feedback Dashboard solution seems to have the potential to be applied in automatized work environments in the factory context. The production metrics that were used, such as the utilization rate of the machine and the recovery time after failures are related to factory work where humans are controlling automated machines. Similar metrics for keeping the process running and solving exceptional situations quickly could be used also in process control type of tasks. Worker Feedback Dashboard was designed so that it is possible to change the production metrics if needed and thus, by redefining the production metrics, the solution could be applied to other kinds of work tasks as well. However, finding suitable metrics to be quantified can be challenging. Although in this study the production metrics were defined in collaboration with the operators, the operators still felt that the metrics did not reflect all relevant aspects of their work objectives. The demands and the nature of work are variable in different factories and not all aspects of work achievements are such that they could be measured. Choosing the production metrics for different jobs should be done carefully and in collaboration with workers and other relevant factory stakeholders. The production metrics should be positive and encouraging, focused on successes and workers' efforts. In addition, the user should be able to select the metrics that are shown in the main view for each work shift, not only the ones that are tracked to follow longer-term evolvement.

The limitations of the field study include the small number of participants and factories involved. Thus, the results can be seen as preliminary and indicative, and not generalizable as such. The participants do not represent all kinds of factory workers, but the sample is particularly interesting, as the participants can be seen as early representatives of the Operator 4.0 vision (Romero et al. 2016), due to the requirements of their work. The study supports the expectation that wearable trackers have potential in supporting the factory workers' occupational health, safety and productivity (Romero et al. 2018).

The study results contribute to the field of Human–Computer Interaction (HCI) by increasing design-relevant understanding of work-related systems and providing understanding of a factory work context, which has received less attention in HCI research. The findings can be used by practitioners to design for quantified worker feedback, also beyond the factory floor context. The results encourage experimentation and trials of solutions, which involve tracking of well-being and work performance or other kind of quantification of a worker. In the future, it would be interesting to see results of testing similar solutions within other work contexts. Furthermore, we encourage the ethical questions related to design, usage trials and adoption of solutions at workplaces to be further addressed and discussed.

9 Conclusions

The results of this field study provide understanding of the experiences, perceived benefits and concerns related to the usage of the Worker Feedback Dashboard solution that offers a new kind of data-driven feedback to factory floor workers on their work achievements and well-being at work. The results indicate that provided feedback may bring benefits related to well-being as well as related to work performance, and thus would support the factory floor workers of the future, Operator 4.0, whose work is changing towards more autonomous, requiring new skills from workers. The results highlight three design implications for quantified worker solutions: presenting meaningful overviews, providing guidance to act based on the feedback and refraining from too pervasive quantification to avoid narrowing down the meaningful aspects in one's work. The results encourage paying particular attention to the ethical questions but also experimenting with similar solutions on the factory floor and other work contexts.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics approval The Ethics Committee of the research organisation conducted an ethical review before the study.

Consent to participate All participants signed a written informed consent form.

References

Ackerman MS (2000) The intellectual challenge of CSCW: the gap between social requirements and technical feasibility. Human-Comput Interact 15(2–3):179–203

Asimakopoulos S, Asimakopoulos G, Spillers F (2017) Motivation and user engagement in fitness tracking: heuristics for mobile healthcare wearables. Informatics 4(1):5

Choe EK, Lee NB, Lee B, Pratt W, Kientz JA (2014) Understanding quantified-selfers' practices in collecting and exploring personal data. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems, pp 1143–1152

Chung C-F, Gorm N, Shklovski I, Munson S (2017) Finding the right fit: understanding health tracking in workplace wellness programs. In: Proceedings of the 2017 CHI conference on human factors in computing systems, ACM, pp 4875–4886

De Zambotti M, Cellini N, Goldstone A, Colrain IM, Baker FC (2019) Wearable sleep technology in clinical and research settings. Med Sci Sports Exerc 51(7):1538

Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango Bo, Buote R et al (2020) Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR mHealth uHealth 8(9):e18694

Gagné M, Deci EL (2005) Self-determination theory and work motivation. J Organ Behav 26(4):331–362

Gavriloff D, Sheaves B, Juss A, Espie CA, Miller CB, Kyle SD (2018) Sham sleep feedback delivered via actigraphy biases daytime symptom reports in people with insomnia: implications for insomnia disorder and wearable devices. J Sleep Res 27(6):e12726

Gorecky D, Schmitt M, Loskyll M, Zühlke D (2014) Human-machine-interaction in the industry 4.0 era. In: Industrial Informatics (INDIN), 2014 12th IEEE International Conference on Industrial Informatics (INDIN), IEEE, pp 289–294

Gouveia R, Karapanos E, Hassenzahl M (2018) Activity tracking in vivo 2018. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, ACM, p 362

Gurrin C, Smeaton AF, Doherty AR (2014) Lifelogging: personal big data. Found Trends Inform Retr 8(1):1–125

Hackman RJ, Oldham GR (1975) Development of the job diagnostic survey. J Appl Psychol 60(2):159

Heikkilä P, Honka A, Kaasinen E (2018a) Quantified factory worker: designing a worker feedback dashboard. In: Proceedings of the 10th Nordic Conference on Human–Computer Interaction, pp 515–523. ACM, 2018

Heikkilä P, Honka A, Mach S, Schmalfuß F, Kaasinen E, Väänänen K (2018b) Quantified factory worker - expert evaluation and ethical considerations of wearable self-tracking devices. In: Proceedings of the 22nd International Academic Mindtrek Conference, ACM, pp 202–211

ISO 9241–210 (2010) Ergonomics of human system interaction-Part 210: human-centred design for interactive systems. International Standardization Organization (ISO). Switzerland

Kaasinen E, Liinasuo M, Schmalfuß F, Koskinen H, Aromaa S, Heikkilä P, Honka A et al (2018) A worker-centric design and evaluation framework for operator 4.0 solutions that support work

well-being. In: IFIP Working Conference on Human Work Interaction Design. Springer, Cham, pp 263–282

Kaasinen E, Roto V, Hakulinen J, Heimonen T, Jokinen JPP, Karvonen H, Keskinen T et al (2015) Defining user experience goals to guide the design of industrial systems. Behav Inform Technol 34(10):976–991

Lavallière M, Burstein AA, Arezes P, Coughlin JF (2016) Tackling the challenges of an aging workforce with the use of wearable technologies and the quantified-self. Dyna 83(197):38–43

Li I, Dey A, Forlizzi J (2010) A stage-based model of personal informatics systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 557–566

Li X, Dunn J, Salins D, Zhou G, Zhou W, Schüssler-Fiorenza SM, Rose DP et al (2017) Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. PLoS Biol 15(1):1–30

Lupton D (2016) The diverse domains of quantified selves: self-tracking modes and dataveillance. Econ Soc 45(1):101–122

Kuhn S, Muller MJ (1993) Participatory design. Commun ACM 36(6):24–29

Masson CB, Martin D, Colombino T, Grasso A (2016) The device is not well designed for me" on the use of activity trackers in the workplace? In: Proceedings of the 12th International Conference on the Design of Cooperative Systems (COOP 2016), Springer, Cham, pp 39–55

MIT Technology Review: Stephen Wolfram on Personal Analytics. Retrieved November 24, 2020 from https://www.technologyreview.com/s/514356/stephen-wolfram-on-personal-analytics/

Moore VP, Piwek L (2016) Regulating wellbeing in the brave new quantified workplace. Empl Relat 39(3):308–316

Pantzar M, Ruckenstein M (2015) The heart of everyday analytics: emotional, material and practical extensions in self-tracking market. Consum Mark Cult 18(1):92–109

Pardamean B, Soeparno H, Budiarto A, Mahesworo B, Baurley J (2020) Quantified self-using consumer wearable device: predicting physical and mental health. Healthc Inform Res 26(2):83–92

Piwek L, Ellis DA, Andrews S, Joinson A (2016) The rise of consumer health wearables: promises and barriers. PLoS Med 13(2):1–9

Rock Health, Digital Health Consumer Adoption Report 2019. Retrieved November 24, 2020 from https://rockhealth.com/reports/digital-health-consumer-adoption-report-2019/

Romero D, Bernus P, Noran O, Stahre J, Fast-Berglund A (2016) The Operator 4.0: human cyber-physical systems & adaptive automation towards human-automation symbiosis work systems. In: IFIP International Conference on Advances in Production Management Systems, Springer, Cham, pp 677–686

Romero D, Mattsson S, Fast-Berglund A, Wuest T, Gorecky D, Stahre J (2018) Digitalizing occupational health, safety and productivity for the operator 4.0. In: IFIP International Conference on Advances in Production Management Systems. Springer, Cham, pp 473–481

Roto V, Kaasinen E, Heimonen T, Karvonen H, Jokinen JPP, Mannonen P, Nousu H et al (2017) Utilizing Experience Goals in Design of Industrial Systems. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, pp 6993–7004

Schuler D, Namioka A (eds) (1993) Participatory design: Principles and practices. CRC Press, Boca Raton

Selke S (ed) (2016) Lifelogging: digital self-tracking and lifelogging-between disruptive technology and cultural transformation. Springer, Berlin

Stiglbauer B, Weber S, Batinic B (2019) Does your health really benefit from using a self-tracking device? Evidence from a longitudinal randomized control trial. Comput Hum Behav 94:131–139

Swan M (2013) The quantified self: fundamental disruption in big data science and biological discovery. Big Data 1(2):85–99

Vanderhaegen F, Wolff M, Mollard R (2020) Non-conscious errors in the control of dynamic events synchronized with heartbeats: a new challenge for human reliability study. Saf Sci 129:104814

Wolf G (2010) The data-driven life. The New York Times, New York

Zimmerman J, Forlizzi J, Evenson S (2007) Research through design as a method for interaction design research in HCI. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 493–502