Mirror, Mirror on the Wall: Exploring Ubiquitous Artifacts for Health Tracking

Jakob Karolus
jakob.karolus@ifi.lmu.de
LMU Munich
Munich, Germany

Eva Braß
eva.brass@gmx.de
LMU Munich
Munich, Germany

Thomas Kosch
kosch@tk.tu-darmstadt.de
TU Darmstadt
Darmstadt, Germany

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

Paweł W. Woźniak
pawel.wozniak@chalmers.se
Chalmers University of Technology
Gothenburg, Sweden

ABSTRACT

While fitness trackers are increasingly popular among users, recent studies have shown that the health benefits of wearing a tracker are not apparent. The need to explicitly retrieve data can lead to limited benefits. Understanding how users can access, understand, and reflect on their data can lead to building systems that benefit our wellbeing. In this work, we explore the feasibility of using ubiquitous artifacts for unobtrusive feedback in health tracking. We evaluated a concept based on design dimensions for personal visualization on a smart mirror in a user study. Our design puts emphasis on the temporality of presented data. Participants found the visualizations comprehensive, rating cardiac and inertial data most useful as well as approved the different levels of temporal aggregation. Our work contributes findings on how to represent health-related data with ubiquitous artifacts to increase users' awareness.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in ubiquitous and mobile computing**; *Empirical studies in interaction design*; *Ubiquitous and mobile devices.*

KEYWORDS

Fitness tracking, smart mirror, health awareness, ubiquitous artifacts

ACM Reference Format:
Jakob Karolus, Eva Braß, Thomas Kosch, Albrecht Schmidt, and Paweł W. Woźniak. 2021. Mirror, Mirror on the Wall: Exploring Ubiquitous Artifacts for Health Tracking. In *20th International Conference on Mobile and Ubiquitous Multimedia (MUM 2021), December 5-8, 2021, Leuven, Belgium.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3490632.3490671

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MUM 2021, December 5-8, 2021, Leuven, Belgium
© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-8643-2/21/05…$15.00
https://doi.org/10.1145/3490632.3490671

1 INTRODUCTION

Fitness trackers have permeated everyday interactions. Last year, 29% of the US population tracked their health and fitness using an app [18]. As tracker technologies mature, studies of their long-term impact become available. Recent results indicate that current trackers do not offer long-term health benefits to the majority of users [25]. As more and more people use trackers, ensuring that their potential for improving the society's wellbeing is used effectively remains a challenge. Additionally, as tracker models possess more and more features and measurement modalities, it remains a question which physiological measures can be meaningful to users and how they are to be presented.

Research in Human-Computer Interaction (HCI) has attempted to understand how users interact with trackers and how one's personal informatics experience changes over time. Several models that emphasize the importance of data choice, initial curiosity and intermittent lapsing were developed [13, 30, 37]. A key aspect of all those models is reflection — the 'holy grail of personal informatics'. Enabling users to make sense of their tracking data to reach meaningful conclusions, which may lead to improvements in wellbeing, is the key design goal of many personal informatics applications. One common issue is that users tend to unintentionally lapse in their tracking routine, hindering reflection [42]. While many explorations have been conducted, it still remains a challenge to understand the design qualities that may foster personal reflection in interactive systems [3]. Here, we propose ubiquitous artifacts — such as a smart mirror — to foster unobtrusive and near subconscious reflection in the home context, anecdotally a key feature of mirrors [9].

To explore the impact of ubiquitous artifacts on a person's fitness tracking routine, we built a smart mirror that visualizes a wide range of tracking parameters for physiological data gathered by an Empatica E4 wristband. Our concept is based on the four design dimensions of personal visualization [24] and additionally puts emphasis on the temporality of the presented data, being a significant factor for reflection [4, 33]. The combination of three distinct visualization addresses the "six kind of questions" as presented by Li et al. [31]. We evaluated the system in a user study, which explored user expectations and perceptions of a smart fitness mirror, while being presented with different temporal aggregation levels of gathered data. Our results show that users appreciated the appearance of a smart mirror, as it not only served as a means to access personal fitness data, but additionally retained its original purpose of being a mirror, hence integrating seamlessly into a person's daily routine. Participants found the visualizations comprehensive and informative. While skin conductance received mixed ratings, cardiac and inertial data was rated most useful.

This work contributes the following: (1) The design and implementation of a smart interactive mirror that visualizes fitness data based on a feature-rich tracker; (2) the evaluation of the system; (3) insights on user preferences of physiological measures to be tracked; and (4) considerations for the design of future ubiquitous artifacts that provide unobtrusive review of personal fitness data.

2 RELATED WORK

Research in personal informatics has grown significantly over the last decade. For an extensive overview of methods and topics in this field, we refer the reader to a survey by Epstein et al. [12], and a more theoretic approach provided by Rapp et al. [39]. To inform our current work, we present related research on understanding fitness tracking, and review past research related to ubiquitous artifacts, namely situated artifacts and smart mirrors. In particular, we want to investigate if users are able to associate health-related data visualization with corresponding fitness states and how the ubiquity of a smart mirror may facilitate this process.

2.1 Understanding Fitness Tracker Data

Research in personal informatics is now going through a decisive phase. While tracking one's activity is becoming commonplace and widely available, we do not fully know what the possible benefits of tracking are and how they can be achieved through the design of interactive systems [34]. Despite the many ways to make trackers more effective in enhancing the users' wellbeing, reflection and sensemaking are recurring themes. In their models of personal informatics, both Li et al. [30] and Epstein et al. [13] identified reflection as a key feature of a meaningful personal informatics experience. Later, Niess and Wozniak [37] concluded that reflection was necessary for the evolution of tracker goals and sustaining the user's engagement in tracking to offer a long-term experience. These theoretical works on tracking postulate more artifact-centered work which aims to embody the qualities needed for effective reflection on fitness tracker data [4]. Yet, a key issue remains: users tend to neglect or even abort their tracking routine [42]. This leads to the questions of how to keep users engaged while providing sufficient information to foster reflection.

Another strain of research explored what qualities an artifact should possess in order to stimulate reflection. Gullotta et al. [20] explored systems that aimed to engage users in personal informatics. They found that the lack of support for complex life settings and the need for extensive maintenance hindered the possibilities for meaningful reflection. In this regard, work by Bentley et al. [6] supported that feedback via natural language can increase self-understanding. Further, Tang et al. [43] emphasized the role of lapses and explicit management of adherence to fitness tracking routines. Similarly, Agapie et al. [1] showed that allowing users to 'cheat' in fitness tracking, i.e. introducing controlled lapses was beneficial to the tracking experience if the users were empowered to transparently manage the process.

These works show that meaningfully exposing the users to fitness tracker data and allowing the users to effectively manage their data is a key design consideration for a successful fitness tracking system. Our work explores themes suggested by this past research as it looks for ways to enable users to manage their fitness tracking process while reducing system maintenance and putting the user in control.

2.2 Situated Artifacts

The question of how to optimally present fitness tracker data to users is a recognized challenge in the HCI field. Research has investigated how interactive artifacts can offer capabilities beyond the traditional format of on-tracker display or mobile application. A number of research efforts explored if users could benefit from tracker data physicalization [23]. Sauvé et al. [41] built *Loop*, a set of tangible loops designed to be placed in frequented locations in homes. *Loop* exemplified how situated artifacts can be used to visualize fitness data. In a similar vein, Khot et al. [26] built material artifacts in the form of plastic jewelry based on exercise data. They observed that users valued a tangible representation of their physical activity. *TickTockRun* [27] was a situated artifact designed to provoke reflection in users through an ambiguous visualization and ambient sound. These past efforts show a growing need to understand how a situated artifact can effectively visualize fitness data to offer benefits to the user. Our work continues these efforts by examining the suitability of ubiquitous artifacts in the form of an interactive mirror. Compared to situated artifacts, the smart mirror still serves its original purpose and blends into the daily routine. Further, we explore a research question beyond past designs — which physiological parameters are most relevant to users to visualize in a ubiquitous artifact.

We endeavor to explore the mirror form as it was previously shown to be an effective design choice in technologies related to wellbeing and persuasive technologies [35, 38]. Notably, *Virtual Aquarium* and *the Mona Lisa Bookshelf* [36] were one of the first systems to present users with a holistic view of their tracked activities in a situated system. These systems exemplified a past trend in HCI research that focused on using slow technology [22] for persuasive purposes. More recent work indicated that the persuasive approach to designing technology for wellbeing was fundamentally problematic [8]. In this work, we revisit the ambient mirror concept and form. In contrast with past mirrors-like systems, e.g., work by Rapp et al. [38] focusing on explicit reflection, we want to leverage a mirror's ambient presence, keeping its original purpose, while exploring design possibilities for health data reflection.

3 RESEARCH QUESTIONS

Our goal is to provide means for ubiquitous interaction with a person's health tracking environment[1]. It is important to deliver information unobtrusively, seamlessly integrating into everyday artifacts. We want to achieve this goal by employing a smart mirror, that serves primarily as a normal mirror and is recognized by users as such, but additionally display health-related information. The question is, whether current fitness visualizations are still applicable

[1]Trackers and included applications.

and if the mirror can serve as a continuous reminder for users. Hence, we formulate the following research questions:

RQ1: How useful is a smart mirror with regard to health tracking? What are the benefits when deploying a smart mirror to display health data that would increase the user's understanding of their physical activity?

RQ2: What are the design requirements for visualizing physiological data on a smart mirror? Design requirements and guidelines for motivating a physically active lifestyle are plentiful [10]. Which strategies are especially fit to be fulfilled by a smart mirror? Which additional requirements need to be considered? Is there a trade-off between ubiquitousness and informativeness?

4 DESIGN

In the following section, we outline identified design requirements and how we evaluated these in our smart mirror prototype.

4.1 Design Dimensions for Ubiquitous Artifacts

Analysis of personal data and appropriate visualizations can offer substantial insights for individuals about themselves. The area of visualizations for personal context and subsequent visual analytics is still a very active research field [24]. Consequently, we draw on the design space of Huang et al. [24] (*Data*, *Insight*, *Context*, *Interaction*) and identify specifications applicable for ubiquitous artifacts, such as a smart mirror, enabling a more holistic perspective on how to visualize personal data, suited for our study design with different personas. Additionally, we explicitly look at *Ubiquitousness* as a design dimension.

Data. Enabling continuous and straightforward access to personal data is vital for user acceptance. This includes — but is not limited to — ease of use and comfort of the recording device [5] as well as recalling one's data [2]. Another point that surfaced in our focus groups was who should be able to gain access to personal data and how to control for it [11]. We addressed this dimension in our evaluation via interviews, asking participants about their impressions of an ubiquitous mirror and possible privacy concerns. Additionally, participants were polled about the general trustworthiness of the system in questionnaires.

Insight. For an engaging experience, the presented data needs to be comprehensible, including its aggregation as well as how processed data is visualized on the display [14]. Especially for complex data, an appropriate level of abstraction has to be chosen. A major part of our evaluation assessed the comprehensibility of shown visualizations using a set of quiz questions.

Context. Though research has looked at what information and how to embed it into situated artifacts, it is unclear how and when to visualize this information to help with correct data interpretation and recall for ubiquitous artifacts. For our smart mirror, we look at the impact of temporality, including live feedback (as endorsed in the focus groups), as well as mid-term (on a daily basis) and long-term (multiple day) feedback for the user [29]. Additionally, we examine the aspect of personalization and how it can be realized in interviews.

Interaction. Interaction is a vital factor when it comes to exploring one's personal data. It is strongly interconnected to how data is visualized, as unsuited visual metaphors might lead to superfluous interaction. In our evaluation, we use a wizard-of-Oz approach to test interaction with the smart mirror and to enable data exploration.

Ubiquitousness. While the design dimension *Context* covers the user's setting and environment with regard to personal visualization, we gauge how and if a ubiquitous artifact can positively affect health tracking, while still retaining its original function. In other words, is a smart mirror with added personal data still perceived primarily as a mirror or has it lost this particular affordance to become mainly a health-tracking artifact? We examine this facet in interviews.

4.2 Sensing Modalities

Tracking devices employ different sensors to detect activity. Most commonly, inertial sensors (e.g. accelerometers) are used to track motion, for example for step counting by detecting walking patterns. Cardiac data is often recorded using a photoplethysmography sensor (PPG) measuring light reflection on the skin [16], that correlates to heartbeats. The average heart rate can be extracted using peak detection and is an indicator for physical activity. Moreover, heart rate variability (HRV) is said to be connected to the autonomic nervous system [21]. A relatively under-explored sensing modalities is electrodermal activity (EDA) measuring the electrical conductance of the skin [16]. It can be used as an indicator of emotional arousal and stress [19].

To measure a broad range of physiological data while keeping the form factor manageable, we opted for the Empathica E4[2]. The wristband has the size of a watch and can measure acceleration, blood volume pulse and related cardiac features, skin temperature, and skin conductance. We used the Empatica E4 to record demo datasets that are to be visualized on the smart mirror. Recordings were collected over an eight-day period for each dataset (three in total).

4.3 Visualization Concept

In the following, we present our visualization concept consisting of three distinct views that are displayed simultaneously on the mirror (cf. Figure 3). We drew from Li et al.'s work [31] on supporting self-reflection with ubiquitous technologies [3]. For our work, we put a special emphasis on the temporality of data by providing a live-data view as well as a mid-term visualization that covered the current day. Finally, a weekly overview provided insights into mid- to long-term data. We used a combination of visual metaphors [15] (*current value view*), abstract and graph visualizations [14] (*daily overview* and *timeline view*).

Timeline View. The *timeline view* emphasizes trends of physical activity by depicting the user's steps taken for the last seven days. It provides the most data history while compromising on data accuracy. A white circle represent the steps that were taken, while a

[2] www.empatica.com/research/e4

[3] The "six kinds of questions" [31]: *Status, History, Goals, Discrepancies, Context*, and *Factors*, that users ask about their data.

dotted circle indicates the user's goal. The selected day for the *daily overview* is highlighted in yellow. The visualization is illustrated in Figure 1. A mouse was used to mimic touch interaction in a wizard-of-Oz manner that enables day selection.

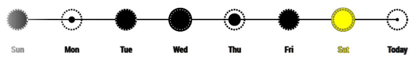

Figure 1: The *timeline view* shows the number of taken steps for each day of the last seven days. The selected day for the *daily overview* is highlighted in yellow. A daily goal of steps is visualized by a dotted circle line on each day. Black and white are inverted for visual clarity.

Daily Overview. The *daily overview* compromises between available data history (current day) and data accuracy (down to sub-hourly values). It visualizes heart rate (red), temperature (green), skin conductance (blue) and inertial data (white) in a radial graph that represents a full day of data as depicted in Figure 2. The view uses an analog clock analogy to mirror recurring times of a day and allows the user to identify phases of physical activity. Through the *timeline view*, the user is able to select individual days and compare the radial graphs.

Current Value View. The *current value view* (cf. Figure 2) view does not provide any information on data history, but rather displays real-time data on heart rate (heart icon), temperature (thermometer icon) and skin conductance (raindrop icon). Colors are chosen in accordance with the *daily overview*. Inertial data is not displayed in this particular view, as metrics focus on its time derivatives. This view provides a direct connection between the user's current physical state and displayed information. Being able to directly identify events of physical activity, e.g. elevated heart rate after an afternoon run, facilitates trust in the sensor data.

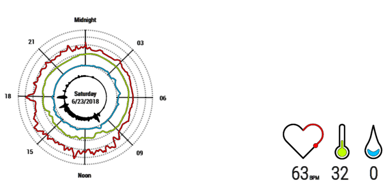

Figure 2: Left: the *daily overview* visualizes the inertial data (white), the skin conductance (blue), the skin temperature (green) and the heart rate (red) over the 24-hour period of one day. Right: the *current value view* shows the current values of the heart rate (heart icon), temperature (thermometer icon) and skin conductance (raindrop icon). Black and white are inverted for visual clarity.

5 EVALUATION

We conducted a user study employing a within-subject design with the *visualization type* as independent variable, consisting of the three levels: *daily overview*, *timeline view*, and *current value view*. All views were presented at the same time during the experiment as depicted in Figure 3. We investigate our design using questionnaires measuring comprehensiveness, usefulness, appearance and informativeness of the visualizations. Furthermore, we addressed interaction, trustworthiness and privacy through user ratings. A short interview concluded the study.

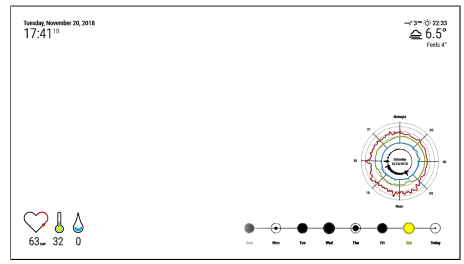

Figure 3: A screen shot of the smart mirror system. The *current value view* is placed in the lower left corner, the *timeline view* and *daily overview* in the lower right corner. As users would display more than just the fitness visualizations, a digital clock and the current weather are added in the upper corners. Black and white are inverted for visual clarity.

5.1 Apparatus

Visualizations were shown on a custom-built smart mirror consisting of a framed one-way mirror in front of a 26 inch LCD monitor, as depicted in Figure 4. Thus, it was possible to use the device as a mirror. Information was displayed on the LCD monitor. A Raspberry Pi 3 Model B+ running the MagicMirror2[4] framework served as the computing unit. Touch interaction with the *timeline view* was implemented by the experimenter in a wizard-of-Oz manner.

Figure 4: Smart mirror prototype with participants exploring the touch interaction.

5.2 Procedure

Firstly, we introduced prospective participants to our study and informed them about the study procedure. After providing consent, the participants were asked to fill out a background questionnaire querying demographics, their experience with fitness tracking and mirror usage as well as self-assess their own fitness. Additionally,

[4]www.magicmirror.builders

we asked them to fill in the *International Physical Activity Questionnaire (IPAQ-SF)*[5], which returns one out of three physical activity categories (low, moderate, and high). This took about 5 to 10 minutes.

After explaining and introducing the participants to each *visualization type*, we tasked them to answer a series of comprehension questions tailored towards the shown visualizations. We created three data sets, that mimic different fitness levels (based on the *IPAQ-SF*[5]), including a very sportive persona, one moderately sportive persona and one persona that was not very active based on the collected data from the Empatica E4. The goal of these questions was, on the one hand, to provide an incentive for the participants to explore the visualizations and, on the other hand, to test the comprehensibility of each *visualization type*. The template set of eight questions was randomized and can be seen in Table 1.

Comprehension questions
At what time on day x was person x the most physically active?
Was person x on day x between xx and xx o'clock more active than between xx and xx o'clock?
Was the heart rate/skin conductance on day x at xx o'clock higher than at xx o'clock?
On how many days did person x achieve their goal of taking x steps?
On what day (except today) did person x recorded the least number of steps?
At what time on day x did person A/B/C probably get up from bed?
Do you consider person A/B/C a sportive person?
What is your current heart rate/temperature/skin conductance?

Table 1: Template of the comprehension questions for each data set.

During the experiment, the participants were invited to explore the visualizations for each persona. When the participants felt ready, they could answer the comprehension questionnaire for the respective persona. The same procedure was applied for the other data sets, taking about 20 to 30 minutes.

Afterwards, we evaluated the usefulness and appearance of every *visualization type* using a post-study questionnaires consisting of a total of seven 5-item Likert-Scale questions and one free text comment field for each *visualization type*[6]. Additionally, we included two 5-item Likert-Scale questions and one free text comment field for possible interaction scenarios and six 5-item Likert-Scale questions for general appearance. Answering the questionnaire took 10-15 minutes and was concluded with a short interview allowing participants to further elaborate on their answers and compare the smart mirror system to the fitness tracking method the participant had been using previously. Interviews took between 5 and 20 minutes and were recorded for later transcription. The total time of the experiment did not exceed 75 minutes and participants were paid an allowance of 15 Euros.

5.3 Participants

We recruited 15 participants ($M = 27.7\,y$, $SD = 7.7\,y$, 8 female) from the [removed for review] through mailing lists. Eleven participants had prior experience with fitness tracking; most using Fitbits (5) and Apple Watch (4). Other used devices include Mio, Kingsky, Polar, Samsung devices and smartphone apps. Out of the tracking participants, the mean tracking duration was $M = 14.7\,months$ ($SD = 13.5\,months$). No participant had used a smart mirror before.

The result of the IPAQ-SF showed that 53% of participants fell into the high category, 40% in the moderate and only one person (7%) in the low category. This overlaps with the self-rated sportiness provided by the participants as depicted in Figure 5.

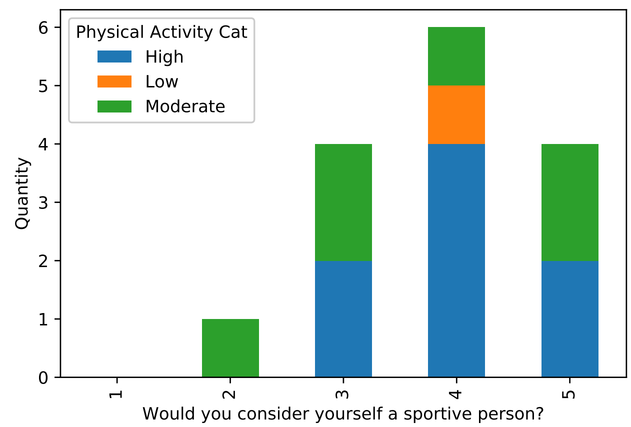

Figure 5: Self-rating of sportiness by the participants in comparison to their physical activity category according to the IPAQ-SF.

6 RESULTS

In this section, we present the aggregated results from the comprehension questions. Lastly, we visualize the post-study questionnaire feedback and perform statistical analysis on the informativeness, usefulness, appearance, and trustworthiness of each *visualization type*.

6.1 Comprehension Questions

On average each participant made two incorrect answers ($min = 1$, $max = 4$), resulting in a success rate of 91.4%. The question with the most incorrect answers (6) was *'At what time on day x was person x the most physically active?'*. Table 2 gives an overview split by persona (A=very sportive, B=moderately sportive, C=not sportive) and visualization. Note that *current value view* was not addressed by any comprehension question.

6.2 Post-Study Questionnaire

All plots are visualized as suggested by Robbins et al. [40] and depict aggregated ratings of questionnaire answers.

6.2.1 Daily Overview. The overall appearance of *daily overview* was rated positively, with most answers leaning to the agreeing and strongly agreeing side (see Figure 6). The visualization was perceived as being informative, while the temperature curve and the skin conductance was rated less useful than both the heart rate curve and accelerometer data.

[5]https://sites.google.com/site/theipaq/ (English, short, self-administered)
[6]six for *timeline view*.

visualization type	Persona	Correct answer ratio	
	A	.96	
daily overview	B	.93	.91
	C	.80	
	A	.90	
timeline view	B	.90	.92
	C	.97	

Table 2: Correct answer ratio split by persona sportiness (A=very, B=moderately, C=not) and *visualization type*. Note that *current value view* was not addressed by any comprehension question.

6.2.2 *Timeline View.* *Timeline view* ratings were positive, both for appearance and informativeness with mostly agreeing and strongly agreeing answers (see Figure 6). Individual visualization aspects such as the time range and the level of aggregation were rated positively, while only two outliers for the step counter visualization and the goal reference line could be registered.

6.2.3 *Current Values View.* Participants' ratings indicate a positive review for the choice of icons in *current value view*, with only two disagreeing answers (see Figure 7). While the visualization was seen as mostly positive, the usefulness of individual sensors varied. Similarly to *daily overview* temperature and skin conductance were rated as less useful than heart rate.

6.2.4 *Interaction.* Participants were instructed to imagine the mirror to be a touch screen to evaluate possible interaction scenarios as depicted in Figure 4. Interaction was then realized in a wizard-of-Oz manner. Ratings were positive except two negative outliers (see Figure 8).

6.2.5 *Overall.* The question *'Do you like the overall appearance of the system'* was rated positively throughout; most answers strongly agreeing (see Figure 7). Privacy-related questions yielded mixed ratings.

Furthermore, we analyzed each *visualization type* with regard to their *informativeness*, *usefulness*, *appearance* and *trustworthiness*. Using a repeated-measures ANOVA with aligned-rank-transformed data as introduced by Wobbrock et al. [44], we identified significant differences among the *visualization type*. We adopted an alpha level of 0.05 for statistical significance testing.

Informativeness. The grand mean of *informativeness* was 4.13 (SD = 0.73). The highest average score ($M = 4.53, SD = 0.64$) was achieved using *current value view*, while the lowest average score ($M = 3.73, SD = 0.59$) occurred for *daily overview*. A repeated measures ANOVA identified a significant main effect for *visualization type*: $F(2, 42) = 5.72, p < .01$. A Tukey post hoc test revealed significant differences between *daily overview* and *current value view* ($p < .01$). No further statistically significant pair-wise differences were found.

Usefulness. The grand mean of *usefulness* was 3.94 ($SD = 0.61$). The highest average score ($M = 4.18, SD = 0.55$) was achieved using *timeline view*, while the lowest average score ($M = 3.82, SD = 0.71$) occurred for *current value view*, *daily overview* recorded a score of $M = 3.83$ ($SD = 0.53$). A repeated measures ANOVA did not reveal a significant effect for *visualization type*.

Appearance. The grand mean of *appearance* was 4.2 ($SD = 0.78$). The highest average score ($M = 4.51, SD = 0.47$) was achieved using *current value view*, while the lowest average score ($M = 3.73, SD = 0.96$) occurred for *daily overview*. We identified a significant main effect for *visualization type*: $F(2, 42) = 3.55, p < .05$. A Tukey post hoc test revealed significant differences between *daily overview* and *current value view* ($p < .05$). No further statistically significant pair-wise differences were found.

Trustworthiness. The grand mean of *trustworthiness* was 4.13 ($SD = 0.63$). The highest average score ($M = 4.33, SD = 0.49$) was achieved using *timeline view*, while the lowest average score ($M = 3.93, SD = 0.70$) occurred for *current value view*, *daily overview* recorded a score of $M = 4.13$ ($SD = 0.64$). A repeated measures ANOVA did not reveal a significant effect for *visualization type*.

6.3 Interviews

Concluding interviews were recorded and transcribed verbatim. We followed the pragmatic approach to thematic analysis as presented by Blandford et al. [7] to conduct a focused analysis. After two researchers coded a representative 20% of the transcribed interviews, an initial coding tree was established through iterative discussion. Afterwards, the rest of the material was split evenly between the two researchers for final coding. The coding tree was finalized in a final discussion, also looking for high-level themes. We have identified two main themes: Temporality and Ubiquitousness. The following section elaborates on these themes.

6.3.1 *Temporality.* Different aspects of Temporality were a recurring point throughout the interviews. Participants highlighted the need for additional interaction with data, providing details when desired:

> *But if you want to know your exact heart rate when doing exercises. You want to know if it was really exhausting or not, just by clicking on it. (P8)*

While *timeline view* provided means for interaction, participants expressed the need for additional interaction with other views to customize their temporal perspective. Yet, interaction focused solely on providing further details for exploration:

> *It would be nice to be able to interact with sections of the daily overview, resulting in the section to be highlighted and be more detailed. (P10)*

Especially being able to interact with historic data in *timeline view* was endorsed and could be further extended:

> *You can see the last seven days to get a better overview. Theoretically, one might add even more days. (P2)*

This also highlights the need for extended control over time, e.g. being able to define periods of aggregation and comparisons to previous periods with regard to different metrics:

> *I can check my data and compare whether I got better or worse [on a conventional tracking app]. The smart mirror only shows me the last seven days. (P14)*

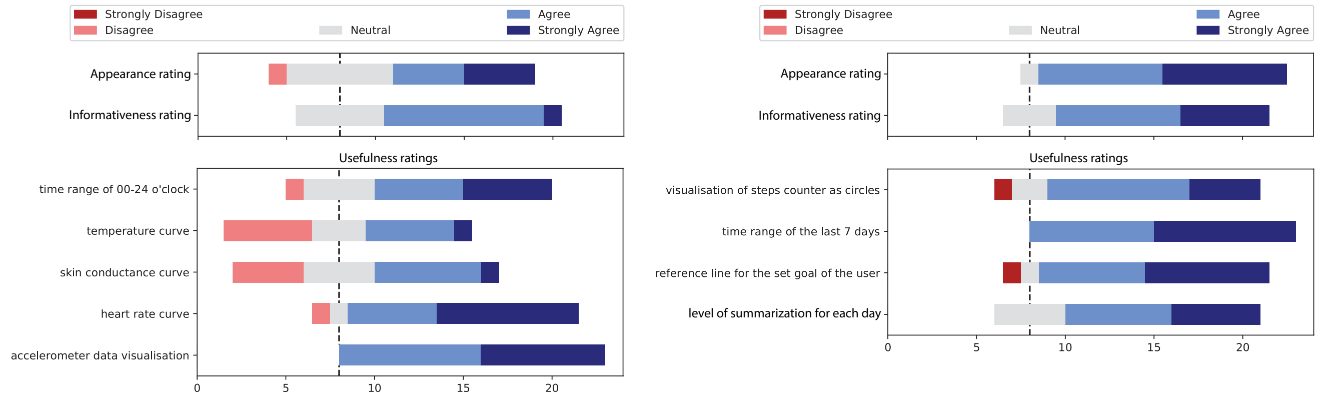

Figure 6: Ratings for *daily overview* (left) and *timeline view* (right).

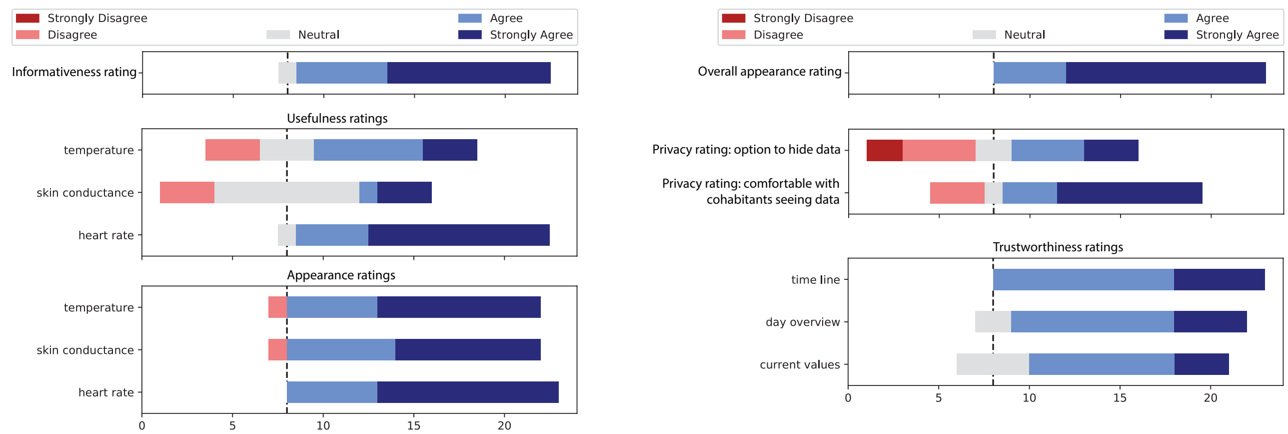

Figure 7: Ratings for *current value view* (left) and ratings for overall appearance, privacy aspects and trustworthiness (right).

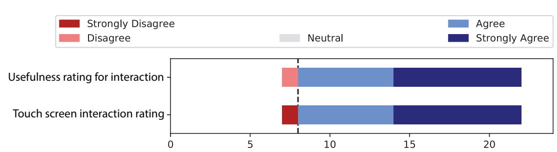

Figure 8: Ratings for possible interaction opportunities.

The smart mirror provided participants with different temporal resolutions of fitness data. The display of live data through *current value view* particularly stimulated their curiosity as it provided immediate and direct feedback that participants could relate to:

> *I found the heart rate interesting; the fact that you always see how your heart beats. "Oh damn, I am excited", now it goes up, "darn it, go back down." (P8)*

To get a general read overview of one's fitness, *daily overview* was endorsed as it provided users with a good overview of important metrics that can be skimmed quickly:

> *I liked the [daily overview] the most. You can get a lot of information on first glance; concise and clear. (P7)*

As part of the daily routine, the mirror also provided feedback on a regular and constant basis:

> *I found this very useful. You can get info on your current state whenever you start your day, like how did you sleep. (P8)*

6.3.2 Ubiquitousness. Another major theme for participants was the idea of having a ubiquitous artifact for health tracking. Users reported that a smart mirror did not require them to explicitly retrieve their fitness status compared to conventional trackers:

> *It is cool that I have the information on the mirror, like as self-reflection, since I look at myself anyway, that would be optimal. And then, you can additional see statistics. A cool concept, very fitting for a mirror. (P7)*

Instead of being perceived as an artifact or system to track one's health, the smart mirror is rather still perceived as a mirror and hence as a part of the furniture, making it aesthetically pleasing:

> *It doesn't bother me, it is still a mirror. (P1)*

This helped participants to identify and understand discrepancies in their fitness routine. In other words, the smart mirror helped them to trigger reflection on a subconscious level:

> *That would be intriguing, cause I personally like it that I keep having that information in my sight. Then, I can tell that I haven't been very active today and the last three days. "How about going outside for a bike ride?" (P15)*

When asked about the possible settings, participants commented on the potential as well as the problematic nature for shared tracking and its seamless integration into current tracking ecosystems:

> *(...) that I can share the mirror with my partner and change between users, e.g. having two different avatars on the bottom to tell whether this is my data or theirs. (P11)*

7 DISCUSSION

The analysis of our comprehension questions and the participants' ratings (post-study questionnaires) show that users of the smart mirror were able to understand and interpret the presented visualization regardless of the *visualization type*. All views were rated positively in terms of usefulness and information content. Importantly, participants expressed high trust in the displayed sensor data, an important criteria for proper self reflection on one's physical activity. This is surprising as past work indicated that users (and especially long-term users) of fitness trackers often doubted the reliability of the data provided [37]. It remains an open question if the alternative form of data presentation may have led to more trust in the sensor data.

We received mixed opinions about the intrusiveness of the smart mirror from questionnaires and interviews. While the opportunity to permanently display fitness data offered convenience, it can be decremental for the user's motivation when being confronted with recent failures in reaching a set goal. On the other hand, reminding the user about successfully reached goals can be a boost for motivation [32]. Hence, we advice to **carefully outweigh positive and negative motivation to personal needs (RQ2)** as a smart mirror is more difficult to avoid.

The smart mirror was perceived as an ambient device that blended in with the home environment providing a seamless reminder of the user's fitness state thus supporting health awareness as indicated by the interviews. While this is one of the biggest advantages compared to fitness tracking apps and websites, it also lacks necessary mobility that the latter provides. A **combined solution that integrates a smart mirror in existing fitness tracking infrastructure (RQ1)** is to be preferred. This also seems feasible as commercial trackers are often offered with additional accessories such as smart scales.

Out of the four provided sensor modalities (inertial, cardiac, skin conductance, temperature), the data from the accelerometer and cardiac data was perceived most useful and informative, while temperature and skin conductance were perceived to be less useful. For the latter, we identified the problem being the missing link between the actual sensor data and its implication for one's health. Yet, participants expressed their interest in this relatively new modality. These results indicate that the increasing number of possible activity metrics in new trackers calls for **extensive user support during the 'choosing metrics' phase (RQ1, RQ2)** [13] of a personal informatics experience.

Furthermore, participants commented on individual, possibly misleading, metrics that are not optimized for their specific kind of physical activity, for example displaying a step counter metric when the user is mainly cycling. Here, we identified the need for **personalization on a modality level** — which sensor data to show — **but also on a temporal aggregation level (RQ2)**, thus allowing to adjust for different fitness behavior and different kind of sports. Even more so, participants remarked that such a mirror should be able to support a **multi-user mode (RQ2)** as well as **be context-aware (RQ2)**, e.g. by providing time-of-day-adjusted data. This also offers a design opportunity for multiple users to gather around a mirror and reflect in a social context, which was identified as desired by past work [27, 28].

With this in mind, we directed questions towards possible interaction opportunities when exploring the data on the smart mirror in our questionnaire and interview. The possibility for touch interaction was appreciated and was mentioned mostly in conjunction with data exploration, e.g. for semantic zoom (details on demand) and highlighting as well as showing trends in one's data. Though some participants remarked that this interaction modality would leave unsightly smudges behind and suggested the use of gestures or eye-tracking for interaction. Based on these reviews, we suggest to **design for short interaction time frames (RQ1, RQ2)**, possibly without the need to touch the mirror. This suggests that mirrors could foster ludic interaction [17].

Our participants described the smart mirror as tidy and aesthetically pleasing, which is also confirmed by the ratings. This is in accordance with important design guidelines aimed at aesthetics [10]. The mirror needs to blend in with the environment, but still needs to fulfill its primary use case, hence **shown visualizations need to be inconspicuous but noticeable (RQ2)**. Our results show that participants found the visualizations ambient enough to not be disturbing.

Another important aspect was the mirror's ability to foster health awareness. We determined that most visualizations and sensor modalities were comprehensible, thus allowing reflection on the user's health and physical activity. We attribute the worse ratings for skin conductance to the unfamiliarity with said modality. This confirms a previous guideline suggesting **individual sensor metrics (RQ1)** that the user identify with [13].

Our mixed results regarding obtrusiveness and privacy-related concerns, reinforce the need for customization. While some participants did not mind that cohabitants were able to see their fitness data, others were not keen on sharing their results as expressed in the interview. This finding highlights the need for a **compromise between privacy and intrusiveness (RQ2)** of a smart mirror. Previously outlined guidelines already take this into account, e.g. by allowing for multiple users and context-aware display of data.

Overall, our work shows that there is unused potential in using ubiquitous artifacts such as smart mirrors to help users benefit more from personal informatics. We observed that the properties of the smart mirror may offer new opportunities for increased engagement, social involvement and ad-hoc reflection. On the other hand, our study also indicated that privacy concerns, tailoring content and limiting intrusiveness remain as design challenges. Finally, we recognize that a possible development of consumer-grade smart mirrors may create threats to user integrity. As a user places a mirror, which is a form of a display, in a prominent place in their home, they could be subjected to product placement or even covert behavioral change strategies when accessing their fitness data. Thus, future development of smart mirrors for fitness must be **accompanied by parallel development of ethical rules for creating content (RQ2)** for them.

8 CONCLUSION

In this paper, we evaluated the suitability of a smart mirror to ubiquitously display fitness data. Contrary to traditional approaches to fitness tracking, we proposed a smart mirror that seamlessly integrates into everyday life as a way for tracking the user's health and fitness.

To this end, we created a visualization concept based on the design space of personal visualization [24] to depict data from different modalities, putting an emphasis on the temporality of data through three distinct views: *daily overview, timeline view*, and *current value view*. We evaluated this concept in a user study through questionnaires and interviews. Results show that participants appreciated the opportunity to ubiquitously display their physical activity on a mirror. Furthermore, the visualizations were comprehensive and easily understood.

We provide guidelines for designing health-aware, ubiquitous systems, such as the need for customization of metrics and display parameters. Moreover, we identified the need for short interaction time frames to explore one's data, preferably avoiding touch, e.g. by combining it with existing devices, such as smartphones. This additionally allows for a seamless transition from a traditional tracking app.

In multi-user settings, a compromise between privacy and ubiquitousness of data has to be made. Context-awareness and personalization can be beneficial in achieving this trade-off. To summarize, our results show that using a smart mirror for ubiquitous visualization of health data is feasible and appreciated by users. We believe that a smart mirror may provide seamless integration into everyday life and complements traditional tracking devices. We hope that our prototype will inspire future work on engaging — meaningfully and ethically — with ubiquitous artifacts for reflection.

ACKNOWLEDGMENTS

This research is supported by the European Union's Horizon 2020 Programme under ERCEA grant no. 683008 AMPLIFY.

REFERENCES

[1] Elena Agapie, Daniel Avrahami, and Jennifer Marlow. 2016. Staying the Course: System-Driven Lapse Management for Supporting Behavior Change. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1072–1083.

[2] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L. Cox. 2018. Flexible and Mindful Self-Tracking: Design Implications from Paper Bullet Journals. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, Article 28, 14 pages. https://doi.org/10.1145/3173574.3173602

[3] Eric P.S. Baumer. 2015. Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 585–594. https://doi.org/10.1145/2702123.2702234

[4] Eric P.S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) *(DIS '14)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/2598510.2598598

[5] Steve Benford, Chris Greenhalgh, Gabriella Giannachi, Brendan Walker, Joe Marshall, and Tom Rodden. 2012. Uncomfortable interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, Austin, Texas, USA, 2005–2014. https://doi.org/10.1145/2207676.2208347

[6] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *ACM Trans. Comput.-Hum. Interact.* 20, 5 (Nov. 2013), 30:1–30:27. https://doi.org/10.1145/2503823

[7] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. Qualitative HCI research: Going behind the scenes. *Synthesis lectures on human-centered informatics* 9, 1 (2016), 1–115.

[8] Hronn Brynjarsdottir, Maria Håkansson, James Pierce, Eric Baumer, Carl DiSalvo, and Phoebe Sengers. 2012. Sustainably Unpersuaded: How Persuasion Narrows Our Vision of Sustainability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. ACM, New York, NY, USA, 947–956. https://doi.org/10.1145/2207676.2208539

[9] Charles S. Carver and Michael F. Scheier. 1978. Self-Focusing Effects of Dispositional Self-Consciousness, Mirror Presence, and Audience Presence. *Journal of Personality and Social Psychology* 36, 3 (1978), 324–332. https://doi.org/10.1037/0022-3514.36.3.324

[10] Sunny Consolvo, David W. McDonald, and James A. Landay. 2009. Theory-driven design strategies for technologies that support behavior change in everyday life. *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09* (2009), 405. https://doi.org/10.1145/1518701.1518766

[11] Daniel A. Epstein, Alan Borning, and James Fogarty. 2013. Fine-Grained Sharing of Sensed Physical Activity: A Value Sensitive Approach. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 489–498. https://doi.org/10.1145/2493432.2493433

[12] Daniel A. Epstein, Clara Caldeira, Mayara Costa Figueiredo, Xi Lu, Lucas M. Silva, Lucretia Williams, Jong Ho Lee, Qingyang Li, Simran Ahuja, Qiuer Chen, Payam Dowlatyari, Craig Hilby, Sazeda Sultana, Elizabeth V. Eikey, and Yunan Chen. 2020. Mapping and Taking Stock of the Personal Informatics Literature. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4 (Dec. 2020), 126:1–126:38. https://doi.org/10.1145/3432231

[13] Daniel A Epstein, An Ping, James Fogarty, and Sean A Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 731–742.

[14] Chloe Fan, Jodi Forlizzi, and Anind K. Dey. 2012. A spark of activity: exploring informative art as visualization for physical activity. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*. Association for Computing Machinery, Pittsburgh, Pennsylvania, 81–84. https://doi.org/10.1145/2370216.2370229

[15] Jon Froehlich, Tawanna Dillahunt, Predrag Klasnja, Jennifer Mankoff, Sunny Consolvo, Beverly Harrison, and James A. Landay. 2009. UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, Boston, MA, USA, 1043–1052. https://doi.org/10.1145/1518701.1518861

[16] Maurizio Garbarino, Matteo Lai, Dan Bender, Rosalind W. Picard, and Simone Tognetti. 2015. Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. *Proceedings of the 2014 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming Healthcare Through Innovations in Mobile and Wireless Technologies", MOBIHEALTH 2014* (2015), 39–42. https://doi.org/10.4108/icst.mobihealth.2014.257418 arXiv:0510181 [math]

[17] William W. Gaver, John Bowers, Andrew Boucher, Hans Gellerson, Sarah Pennington, Albrecht Schmidt, Anthony Steed, Nicholas Villars, and Brendan Walker. 2004. The Drift Table: Designing for Ludic Engagement. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (Vienna, Austria) *(CHI EA '04)*. ACM, New York, NY, USA, 885–900. https://doi.org/10.1145/985921.985947

[18] gfk. 2018. Global Studies – Fitness tracking. http://www.gfk.com/global-studies/global-studies-fitness-tracking/

[19] Martin Gjoreski, Hristijan Gjoreski, Mitja Luštrek, and Matjaž Gams. 2016. Continuous Stress Detection Using a Wrist Device: In Laboratory and Real Life. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (Heidelberg, Germany) *(UbiComp '16)*. ACM, New York, NY, USA, 1185–1193. https://doi.org/10.1145/2968219.2968306

[20] Rebecca Gulotta, Jodi Forlizzi, Rayoung Yang, and Mark Wah Newman. 2016. Fostering Engagement with Personal Informatics Systems. In *Proceedings of the*

2016 ACM Conference on Designing Interactive Systems (Brisbane, QLD, Australia) *(DIS '16)*. ACM, New York, NY, USA, 286–300. https://doi.org/10.1145/2901790.2901803

[21] S Guzzetti. 2014. *Heart rate variability*. Vol. 2. 450–454 pages. https://doi.org/10.1556/APhysiol.88.2001.3-4.4

[22] Lars Hallnäs and Johan Redström. 2001. Slow Technology – Designing for Reflection. *Personal Ubiquitous Comput.* 5, 3 (Jan. 2001), 201–212. https://doi.org/10.1007/PL00000019

[23] Trevor Hogan, Uta Hinrichs, and Eva Hornecker. 2017. The Visual and Beyond: Characterizing Experiences with Auditory, Haptic and Visual Data Representations. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. ACM, New York, NY, USA, 797–809. https://doi.org/10.1145/3064663.3064702

[24] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. 2015. Personal Visualization and Personal Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 21, 3 (March 2015), 420–433. https://doi.org/10.1109/TVCG.2014.2359887

[25] John M Jakicic, Kelliann K Davis, Renee J Rogers, Wendy C King, Marsha D Marcus, Diane Helsel, Amy D Rickman, Abdus S Wahed, and Steven H Belle. 2016. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. *Jama* 316, 11 (2016), 1161–1171.

[26] Rohit Ashok Khot, Larissa Hjorth, and Florian'Floyd' Mueller. 2014. Understanding physical activity through 3D printed material artifacts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3835–3844.

[27] Kristina Knaving and PawełWózniak. 2016. TickTockRun: Towards Enhancing Communication in Runner Families. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion* (San Francisco, California, USA) *(CSCW '16 Companion)*. ACM, New York, NY, USA, 309–312. https://doi.org/10.1145/2818052.2869114

[28] Kristina Knaving, PawełWoźniak, Morten Fjeld, and Staffan Björk. 2015. Flow is Not Enough: Understanding the Needs of Advanced Amateur Runners to Design Motivation Technology. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 2013–2022. https://doi.org/10.1145/2702123.2702542

[29] Moon-Hwan Lee, Seijin Cha, and Tek-Jin Nam. 2015. Patina Engraver: Visualizing Activity Logs as Patina in Fashionable Trackers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1173–1182. https://doi.org/10.1145/2702123.2702213

[30] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 557–566.

[31] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing (UbiComp '11)*. Association for Computing Machinery, Beijing, China, 405–414. https://doi.org/10.1145/2030112.2030166

[32] Edwin A Locke and Gary P Latham. 2013. *New developments in goal setting and task performance*. Routledge.

[33] Sus Lundgren. 2013. Toying with Time: Considering Temporal Themes in Interactive Artifacts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 1639–1648. https://doi.org/10.1145/2470654.2466217

[34] Sean A. Munson. 2017. Rethinking Assumptions in the Design of Health and Wellness Tracking Tools. *Interactions* 25, 1 (Dec. 2017), 62–65. https://doi.org/10.1145/3168738

[35] Tatsuo Nakajima and Vili Lehdonvirta. 2013. Designing Motivation Using Persuasive Ambient Mirrors. *Personal Ubiquitous Comput.* 17, 1 (Jan. 2013), 107–126. https://doi.org/10.1007/s00779-011-0469-y

[36] Tatsuo Nakajima, Vili Lehdonvirta, Eiji Tokunaga, and Hiroaki Kimura. 2008. Reflecting Human Behavior to Motivate Desirable Lifestyle. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems* (Cape Town, South Africa) *(DIS '08)*. ACM, New York, NY, USA, 405–414. https://doi.org/10.1145/1394445.1394489

[37] Jasmin Niess and Paweı W Wozniak. 2018. Supporting Meaningful Personal Fitness: the Tracker Goal Evolution Model. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 171.

[38] Amon Rapp, Alessandro Marcengo, Luca Buriano, Giancarlo Ruffo, Mirko Lai, and Federica Cena. 2018. Designing a Personal Informatics System for Users without Experience in Self-Tracking: A Case Study. *Behaviour & Information Technology* 37, 4 (April 2018), 335–366. https://doi.org/10.1080/0144929X.2018.1436592

[39] Amon Rapp and Maurizio Tirassa. 2017. Know Thyself: A Theory of the Self for Personal Informatics. *Hum.-Comput. Interact.* 32, 5-6 (Nov. 2017), 335–380. https://doi.org/10.1080/07370024.2017.1285704

[40] Naomi B. Robbins, Richard M. Heiberger, Christine Court, and Speakman Hall. 2011. Plotting Likert and Other Rating Scales. *Joint Statistical Meetings* October 2005 (2011), 1058–1066.

[41] Kim Sauvé, Steven Houben, Nicolai Marquardt, Saskia Bakker, Bart Hengeveld, Sarah Gallacher, and Yvonne Rogers. 2017. LOOP: A physical artifact to facilitate seamless interaction with personal data in everyday life. In *Proceedings of the 2017 ACM Conference Companion Publication on Designing Interactive Systems*. ACM, 285–288.

[42] Lie Ming Tang, Jochen Meyer, Daniel A. Epstein, Kevin Bragg, Lina Engelen, Adrian Bauman, and Judy Kay. 2018. Defining Adherence: Making Sense of Physical Activity Tracker Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 37 (March 2018), 22 pages. https://doi.org/10.1145/3191769

[43] Lie Ming Tang, Jochen Meyer, Daniel A. Epstein, Kevin Bragg, Lina Engelen, Adrian Bauman, and Judy Kay. 2018. Defining Adherence: Making Sense of Physical Activity Tracker Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 37 (March 2018), 22 pages. https://doi.org/10.1145/3191769

[44] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963