# Exploring Fitness Tracker Visualisations to Avoid Rumination

Jasmin Niess
University of Luxembourg
Esch-sur-Alzette, Luxembourg
jasmin.niess@gmail.com

Kristina Knaving
RISE Research Institutes of Sweden
Gothenburg, Sweden
kristina.knaving@ri.se

Alina Kolb
LMU Munich
Munich, Germany
alinakolb@gmx.de

Paweł W. Woźniak
Utrecht University
Utrecht, the Netherlands
p.w.wozniak@uu.nl

## ABSTRACT

Fitness trackers encourage users to set goals to improve personal wellbeing, but these goals sometimes remain unmet. Understanding how improved ways of communicating failure to meet fitness goals could help prevent negative thought cycles (rumination) and avoid reduced motivation for physical activity. To address this challenge, we studied how unmet goals can be presented in apps. We designed prototypes that showed unmet fitness goals. Radial and regular bar charts, single-coloured and multicoloured were used in the study. In a survey ($n$ = 165), we compared the four versions and a textual description of the unmet goals. Then, we conducted follow-up interviews ($n$ = 20) to gain a detailed understanding of the perceptions and feelings evoked by the prototypes. We found that bar graphs offered a significantly better potential for reflection and multicoloured charts triggered significantly more rumination. We contribute in-depth insights into designing systems that use goals and avoid potential negative effects of personal tracking.

## CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

## KEYWORDS

Fitness tracking; personal informatics; reflection; rumination; interactive visualisation

**ACM Reference Format:**
Jasmin Niess, Kristina Knaving, Alina Kolb, and Paweł W. Woźniak. 2020. Exploring Fitness Tracker Visualisations to Avoid Rumination. In *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '20), October 5–8, 2020, Oldenburg, Germany*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3379503.3405662

## 1 INTRODUCTION

Personal informatics systems such as mobile fitness apps and fitness trackers, which support users to reflect about their data and monitor their behavior, have become ubiquitous [25]. The commercial proliferation of fitness technologies is reflected in research in personal informatics in the field of Human-Computer Interaction (HCI). However, as previous studies have shown, designing meaningful and supportive personal informatics experiences remains a challenge [27]. Epstein et al. found that, due to certain barriers (e.g. supporting reflection [38]) users lapse throughout their tracking process [20]. In line with that, Niess and Woźniak pointed out in their *Tracker Goal Evolution Model* that fitness tracker users often struggled to connect their numeric tracker goals to their qualitative life goals [43]. Consequently, despite the fact that more and more users wear a fitness tracker every day, we still do not know how to use them to profoundly improve people's lives. Furthermore, it is yet to be determined how to mitigate negative effects of fitness tracker use and failure to reach fitness goals such as declined motivation or negative thought cycles (i.e. rumination).

One of the most extensively studied strategies in the area of personal informatics to help users on their way to wellbeing and foster engagement is goal setting [15]. Almost all commercially available fitness and activity tracking devices are designed to support goals and goal fulfilment [26]. Studies in Psychology also confirm that goals are an effective strategy for fostering positive change, with one such study by Lacroix et al. showing that difficult goals lead to higher levels of performance than easier goals [37]. These findings show opportunities for fitness tracker design. However, when fitness apps and trackers make extensive use of goals, there will always be a fraction of users who do not reach the goals [34, 43], and thus the question of how to communicate failure without demotivating users emerges. Understanding how negative thought cycles (i.e. *Rumination*) can be prevented is a challenge, with recent work adding to that challenge by advocating designing for diversity in personal informatics and avoiding negative effects for diverse user groups [51]. In this work, we investigate potential positive and negative effects of the communication of failure in personal informatics visualisations with a focus on reflection and rumination.

To that end, we studied different visualisations that communicate unmet fitness tracker goals. We conducted a between-subject online study which evaluated prototype visualisations or a textual description of unmet fitness tracker goals. We explored how these visualisations were perceived and if they exhibited potential for users to reflect upon or ruminate over their data. We then conducted follow-up interviews to build a detailed, qualitative understanding of the ways to communicate failure for fitness tracker goals.

*MobileHCI '20, October 5–8, 2020, Oldenburg, Germany*

This paper contributes the following: (1) a between-subject vignette study for visualising unmet goals in physical activity trackers; (2) follow-up interviews with users potentially experiencing goal failure; (3) to the best of our knowledge, the first empirical exploration that showcases how different visualisations of unmet goals affect reflection and rumination; and (4) implications for designing future personal informatics systems to avoid potential negative effects of tracking.

In the remainder of this paper, we first introduce the reader to reflection and rumination literature in Psychology that inspired our inquiry. We then review related work in HCI on personal informatics with a specific focus on designs that foster reflection. This is followed by a discussion of related work on visualisation, as well as literature on motivational design strategies for mobile fitness tracking applications. We then report on the details of our study, including the design of the visualisations. This is followed by the discussion and interpretation of the study results. Finally, we discuss how our results inform the design of future fitness tracker experiences.

## 2 RELATED WORK

Our work is inspired by previous research on reflection and rumination in Psychology. Trapnell and Campbell [54] stated that reflection and rumination are distinct but related concepts. Both concepts describe a process where the attention is focused inwards, upon oneself [41]. Reflection has a positive connotation associated with it [54]: It is motivated through a curiosity about oneself and usually accompanied by positive emotions. If short-term negative emotions occur, they are assumed to motivate exploring oneself in more depth and leading to self-development [58]. In Psychology, different conceptualisations of rumination can be found (for an overview see [49]), however, there are some areas of agreement. Rumination is conventionally defined as negative, repetitive thoughts focusing on oneself, circling around personal failure or loss [41, 49]. Since reflection is seen as a crucial part of the fitness tracking process with the aim to generate self-knowledge and support users on their way to self-improvement [7], the question remains how this knowledge can enrich personal informatics research. In our work, we aim to gain a deeper understanding of how unmet personal tracking goals might inadvertently trigger rumination, or, conversely, support reflection.

### 2.1 Reflection in Personal Informatics

Reflection is considered as a catalyst for positive change [47] in multiple application areas of personal informatics [20]. It is one of the main components of a successful personal informatics process [38, 43]. Even though the concept is widely applied in personal informatics systems, there is no shared understanding of how reflection is defined within the HCI community [6]. Notably, researchers strive to build a holistic understanding of the personal informatics experience and what role reflection plays in it [20, 38, 43]. Previous work showed the potential positive effects of designing for reflection and goal-setting in personal informatics. But potential adversarial effects of fitness tracking are still underexplored [51]. In our work, we aim to address this gap. We explore the opportunities for reflection of fitness tracker visualisations but additionally aim to shed some light on potential negative effects (such as rumination) of displaying unmet fitness tracker goals.

Multiple systems endeavoured to integrate reflection into their designs. We outline some notable examples below. Morrison and Bakayov presented a social fitness activity system that encouraged face-to-face encounters with other users in which they discussed and reflected on their physical exercise [42]. Khot et al. used chocolate treats [35] and 3D-printed personalised artefacts [36] to offer playful reflections on physical activity. Through social interactions [59] and tangible artefacts, these examples had a positive effect on user reflection. However, in everyday life fitness trackers and fitness apps are often used alone, i.e. not in social groups, and users do not include any other artefacts apart from their app or tracker in their Personal Informatics process. Hence, we aim to explore how the interaction with common fitness tracker visualisations affects potential reflection and rumination processes.

### 2.2 Goal Failure

Previous work in Psychology has shown that negative feedback can potentially undermine people's confidence in their ability to pursue their goals and their expectations of success [48]. On a similar note, Aarts et al. [1] found that humans may cease their goal pursuits when their goals are regularly followed by cues or events evoking negative affect, such as highlighting failure to achieve their goal. Hence, feedback can support goal pursuit only if presented meaningfully [39]. However, there is no universal answer for this challenge [23] and presenting goal failure still poses problems in research and practice. Personal informatics research has explored the challenge of managing fitness tracker goals from a variety of angles.

Past research found that fitness tracker goals evolve [43] and negative feelings may be associated with changing a goal [30]. Gulotta et al. [30] identified a number of problems that hinder engagement with achievement-based personal informatics systems. They emphasised avoiding the feeling of failure as important to foster reflection [30]. With our work we strive to extend the insights from Gulotta et al. [30]. We explore how current fitness tracker visualisations might foster reflection and if they potentially trigger rumination.

Epstein et al. [20] observed that it was highly likely that users would lapse at some point of their personal informatics process. While lapses may not have negative consequences per se, research suggests that they should be managed by personal informatics systems. To that end, Agapie et al. [2] endeavored to support users to integrate their lapses into their tracking process in a positive way. Clawson et al. [12] studied craigslist users disposing of tracking technology to find that they could be equally motivated by perceived failure, success or social pressure. On another note, Tang et al. [52] examined reasons behind users not wearing their fitness tracker all the time. The authors emphasised the need to account for these differences in usage, defining adherence as a measure of data completeness in order to provide the means to study fitness tracking experiences in a meaningful way [52]. The works above show that lapses are an inherent part of the fitness tracking experience. As lapses often lead to goal failure, our work explores the ways to communicate unmet fitness tracker goals and how this

affects reflection and rumination in users. Thus, in turn, could help mitigate negative effects of lapses.

### 2.3 Visualisation in Personal Informatics

Data visualisation, if well designed, effectively supports human understanding of quantified data [57]. Personal informatics data, particularly measurements collected through fitness trackers, constitute a case where users desire to gain insight into their everyday activity through visualisation. Thus, fitness tracking apps prominently feature different visualisations. Bar charts and radial bar charts are often used[1]. The bar chart maps quantity through bars whose height is measured by position on a vertical scale, and is one of the oldest and most common chart types, considered to have been invented by William Playfair in 1786 [46]. Radial bar charts are also often used by fitness tracker applications, particularly on smartwatch faces[2]. These are also known as concentric bar charts, circular bar charts/graphs, or (radial) pie gauge charts, are fairly new and there is still some confusion on naming: the name is used to denote either bars plotted on a polar coordinate system, or bars extending from a circular floor. We have chosen to use the name 'radial bar chart' in this paper as relevant related research also terms them as such [9, 56].

Cleveland and McGill [13] found that 'position along a common scale' was the easiest of the most common perception tasks that also included curvature and angle. Furthermore, Croxton [17] found that compared to circles, cubes and squares, bars were more effective for comparing values. Furthermore, Blaschek et al [9] performed tests on small screens and noted that radial bar charts performed significantly worse than bar charts for quickly conveying comparisons at a glance. Epstein et al. [19] explored different visualisations of location and activity data to support users in making sense of their tracking data. The combined location and activity data was presented in a variety of visualisations such as tables, graphs or maps, among others. In their longitudinal qualitative study the authors found that participants preferred tables and graphs for finding patterns in their data.

Moreover, scholars have explored a variety of visualisations to foster reflection and behaviour change. For instance, UbiFit provided feedback for participants regarding their physical activity behaviour on a mobile display [14]. The feedback was presented in two versions; one condition displaying the physical activity progress on the mobile phone background screen (i.e. a *glanceable display*), the other not utilising a glanceable display. The results of the longitudinal, qualitative study showed that participants appreciated the glanceable display and maintained their activity level, but the activity level of the participants in the condition without the glanceable display dropped significantly. This study showcases the potential impact of subtle differences in visualisation on user behaviour. This, in turn, was a contributing factor in our decision to explore designs that might trigger negative thought cycles in an empirical, hypothetical cross-sectional study. In line with the findings from the UbiFit study, Gouveia et al. [28] also showed the potential of glanceable feedback. Their results show the importance of taking subtle differences in the visualisation design into account, since these can potentially lead to behavioural differences (e.g. walking more frequently versus reaching one's step goal). We aim to explore the effect of such subtle differences (e.g. single-coloured versus multicoloured visualisations) in our study.

[1]e.g. by Fitbit, https://www.fitbit.com/gb/app
[2]Most notably Apple Health, https://www.apple.com/apple-watch-series-5/health/

## 3 RESEARCH QUESTIONS

Our overarching goal is to learn more about how to avoid negative effects while tracking. We endeavour to understand how unmet fitness tracker goals can be communicated to users without triggering negative thought cycles and foster reflection. This leads to the following research questions:

- **RQ1**: How can unmet fitness tracker goals be presented with visualisations currently used in typical fitness trackers to avoid negative thought cycles in users?
- **RQ2**: How can unmet fitness tracker goals be presented with visualisations currently used in typical fitness trackers to foster reflection in users?

## 4 VISUALISATION DESIGN

To explore how the communication of failed goals in fitness trackers may trigger reflection and rumination, we designed four visualisations based on commercial fitness tracking products. This work is, to the best of our knowledge, the first study that explores rumination in the context of personal informatics. We therefore engage in a careful exploration given the potential negative impact on users. Consequently, our inquiry focuses on understanding current lived experience. To that end, we used two designs that exemplify the most common features of current fitness tracker visualisations with a focus on step data.

We decided to explore visualisations that inspired existing commercial products, rather than designing our own. In line with [Olson and Kellogg], we believe that building an understanding of systems currently used in commercial products can foster informed future design decisions, and also support reflection on past design. Huang et al. [33] explored visualisations on the meta level. They developed a taxonomy of design dimensions of personal visualisations and personal visual analytics, and called for more research to investigate how the visualisations that 'are already out there and actively used' may impact people's lives and how they can be improved. Our work follows their call and aims to explore designs that resemble commercial products to trigger reflection and rumination.

We designed two different graph types, a radial bar chart and a bar chart, in both single-coloured and multicoloured versions. We are aware that these visualisations are not similar regarding the information they provide, as ordinary bar charts are generally considered best practice when visualising connecting values to categories. This makes it easier to identify if a goal was exceeded in the bar chart, as shown by research on glanceability in smartwatch visualisations [9]. Furthermore, we are aware that the presence of unmet goals is particularly emphasised by the colour red the multicoloured version, while in single-coloured visualisations they are not. However, we still opted for designs that are close to existing commercial products (e.g. multicoloured and single-coloured visualisations in two different graph types) because of the importance to build an understanding of systems currently used in regards to

aspects such as potential negative thought cycles (i.e. rumination). The four graphs are presented in Figure 1.

We chose the bar chart and the radial bar chart because they are both popular in step counting applications and commonly used in tracking devices such as Fitbit, Apple Health and Samsung Gear Fit. However, while the bar chart is arguably one of the most frequent graph types users encounter, the radial bar chart presents a less obviously understood form for the data. The bars in the single-coloured chart are greyish blue, while the multicoloured version colours success with a green hue and failure in the signal hues orange and red. Colours are pre-attentively processed, so have a high chance of being isolated for conscious attention [57]. Both colourings are similar to those of visualisations in commercially available fitness tracker apps. If there is a colour scheme in real examples of tracker visualisations it is often related to whether the goals are met, which is a type of dual encoding common within dashboards in general [22]. We believe that our design choices influence how users interpret their data. According to literature, we expected that the radial bar chart would be less familiar and also more difficult to interpret and compare values correctly, thus making it more open to the user's interpretation. We also predicted that colouring by failure and success would have an effect on how users interpreted the weekly result.

Furthermore, we designed the values visualised—the step counts for each day—to look random, but with a number of days almost reaching the goal. Given that there seems to be little evidence for a very specific step count [55], the exact number of steps itself is arbitrary and unlikely to be important to reach except for psychological reasons. However, failing even by a very low number is sometimes shown as failure in commercial apps — through, for example, colouring. We chose an alternative step goal value—8000 steps—rather than the popular 10000. This is in line with a usage scenario where the users set their own goal, and also closer to estimated step counts of a non-sedentary person [55]. The four screens have been carefully designed to have the same information apart from colour and visualisation type.

## 5 METHOD

This study explores how different ways to visualise unmet fitness goals affect reflection and rumination of users. To that end, we conducted two studies. Study 1 was a between-subject online vignette study with five conditions—four visual representations and a textual description of unmet goals. Study 2 was an interview study with the aim of gaining in-depth insights regarding the user's perception of the visualisations. Furthermore, the qualitative data reported aims to add elements of the lived experience, even if hypothetical. Both studies used the same visualisation prototypes, shown in Figure 1.

## 6 STUDY 1: EXPERIMENTAL VIGNETTE STUDY

First, we conducted a *experimental vignette study* (i.e. a study where we ask participants to see the world through the eyes of a hypothetical person in a specific scenario; a method very similar to tailored scenarios [16]). We chose this method because it offers the means to balance the benefits of experimental research with high internal validity and the advantages of applied research with high external validity [3]. Studies focusing on rumination from Psychology have also opted for experimental vignette studies to explore rumination [10, 53]. We decided to use a vignette as a first step, because presenting participants with designs that might trigger rumination and thus risk their wellbeing was neither in line with our ethical standards as researchers nor with the Declaration of Helsinki [5]. Hence, we decided to conduct a hypothetical study, meaning that we inquired about participants' perceptions of different visualisations or a textual description of unmet fitness goals rather than their own fitness performance. Our study is, to the best of our knowledge, the first empirical exploration focusing on the potential of such visualisations to prompt rumination and reflection. Consequently, we used an experimental approach that offered the highest validity while ensuring no negative influence on the wellbeing of the participants. Experimental vignette studies have been used in a variety of research contexts and it has been shown that they deliver results comparable to 'real-life behaviour' (e.g. [21, 31]).

Before the study, participants were presented with a page informing them that the study was fully anonymous and asking for informed consent. The studies were conducted according to the ethics standards of the conducting institution. According to these rules, the survey was not subject to review by an ethics board.

### 6.1 Participants

The participants in our study 1 were recruited through Amazon Mechanical Turk (MTurk). Theoretical as well as empirical research has pointed out potential benefits of MTurk [40, 45]. Furthermore, previous work assessed participants recruited through MTurk as suitable for visualisation [32] and personal informatics [2, 20] research.

We recruited $n = 165$ participants, aged 18–86, $M = 34.59$, $SD = 9.46$ using Amazon Mechanical Turk (MTurk). Sixty-six participants were male and 99 female. We aimed to recruit a diverse group of participants. A Shapiro-Wilk test revealed that the ages were normally distributed and a Pearson's Product-Moment test revealed no correlation between age and any of the measures. The sample size used was calculated with G*Power with target Cohen's $f = 0.25$. The participants were uniformly distributed among conditions and resided in the United States or the European Union. We required participants who had completed at least 1,000 Human Intelligence Tasks (HITs) (i.e. single tasks such as 'Identify the colour of the object in the picture'), presented on MTurk, with a 95% acceptance rate, in line with past work in HCI [20]. The survey took an average of 6min 21s to complete and the participants received $1.00 as compensation.

### 6.2 Conditions

We designed four different visualisations with two different graph types and two different colour versions. We also included a textual description as a fifth condition. Each condition presented the same weekly goal progress towards a specific step goal. Since there may be a possible interaction effect between feedback and reporting, we did not include additional feedback but solely focused on the representation of the weekly step goal progress. This resulted in five experimental conditions for our study:


**(a) Radial bar chart, multi-coloured, RColour** **(b) Bar chart, multicoloured, BColour** **(c) Radial bar chart, single-coloured, RMono** **(d) Bar chart, single-coloured, BMono**

**Figure 1: The four visualisations designed for the purposes of our study.**

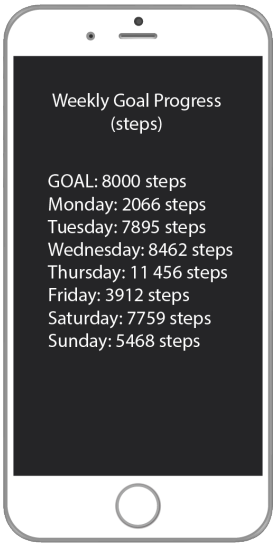


**Figure 2: The text feedback condition — no visualisation, NV used in our study. The text feedback used the same visual style as the other conditions.**

- *Radial bar chart, multicoloured*, RColour, presented in Figure 1a
- *Bar chart, multicoloured*, BColour, presented in Figure 1b
- *Radial bar chart, single-coloured*, RMono, presented in Figure 1c
- *Bar chart, single-coloured*, BMono, presented in Figure 1d
- *Textual description, no visualisation, NV*, presented in Figure 2

### 6.3 Measures

After a brief survey introduction, we asked the participants for their demographic data (age, gender, current occupation, country of birth, country of residence, highest finished education). Afterwards, they were randomly assigned to a condition and presented with the corresponding visualisation. Please note that the method used was adapted from psychology studies [11, 29] and the referenced scales included explanatory text. The visualisation was visible at the top of the screen for the entire duration of the survey. Then, we presented the participants with an open text field where we asked them how looking at the display made them feel. Next, we administered the following measures:

- Rumination Scale, 5 items on a 5-point Likert scale, see Table 1.
- Reflection Scale, 5 items on a 5-point Likert scale, see Table 2.
- Trait rumination scale, 12 items on a 5-point Likert scale.
- Trait reflection scale, 12 items on a 5-point Likert scale.

The rumination scale and the reflection scale were modified versions of the rumination-reflection questionnaire (RRQ) [54] and presented in randomised order. We used the reflection scale and the rumination scale to assess if the way the unmet goal was communicated triggered rumination or reflection. We also inquired about the participants' trait rumination and trait reflection, using the original rumination-reflection questionnaire (RRQ) from Trapnell and Campbell [54]. Trait rumination and trait reflection have previously been identified as personality traits independent of a specific situational context. This means that rumination and reflection manifest themselves as a personality trait (i.e. a somewhat stable, enduring characteristic) and as a state (i.e. a temporary way of being). Hence, in order to explore a user's state rumination and reflection in a methodologically sound way, we also inquired about their trait rumination and reflection, to use it as a covariate in our analysis. Thus, we administered the RRQ to account for these individual differences regarding trait rumination and trait reflection of the participants. This enabled us to separate a participant's general inclination to reflect or ruminate about fitness tracker data from the potential of the different visualisations to activate rumination or reflection.

## 7 STUDY 1: RESULTS

### 7.1 Rumination

We conducted a one-way ANCOVA to investigate the effect of the type of visualisation used on the rumination scale score, controlling for the participants' rumination trait. We found a significant effect, $F_{4,157} = 2.78$, $p < .05$. Post hoc test with Tukey's HSD revealed significant differences between the condition pair BColour — BMono, with $p < .05$. Figure 3 shows the results. There was no significant effect of the covariate, rumination trait scale score, $p > .05$.

| |
|---|
| (1) This app would tend to make me 'ruminate' or dwell over my fitness tracker data for a really long time afterwards. |
| (2) This screen would not make me waste time rethinking training sessions that are over and done with. (R) |
| (3) When I look at this app I feel I would spend a great deal of time thinking back over my embarrassing or disappointing workout moments. |
| (4) This app would make it easy for me to put unwanted thoughts about my fitness tracker data out of my mind. (R) |
| (5) This app would make me focus on aspects of my fitness tracker data I wish I'd stop thinking about. |

**Table 1: The Rumination Scale used in our studies, adapted from work by Trapnell and Campbell [54]. Items were scored on a 5-point Likert scale and items marked with (R) were scored inversely.**

| |
|---|
| (1) I would like to explore my fitness tracker data more with this app. |
| (2) I would often look at this app and analyse my data with it. |
| (3) I would like to analyse why and how I work out with the help of this app. |
| (4) Analysing my fitness tracking data with the help of this app doesn't appeal to me that much. (R) |
| (5) Contemplating about my fitness tracker data with the help of this app isn't my idea of fun. (R) |

**Table 2: The Reflection Scale used in our studies, adapted from work by Trapnell and Campbell [54]. Items were scored on a 5-point Likert scale and items marked with (R) were scored inversely.**

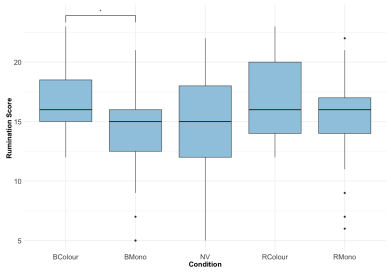


**Figure 3: Rumination scores for the five experimental conditions.**

Since the results presented above found a significant difference between the single-coloured and the multicoloured bar chart (see Figure 3), we analysed the data further and investigated the effect of the colouring scheme (multicoloured, single-coloured) and shape used (radial bar chart, bar chart) on the rumination score. A two-way ANCOVA, controlling for the participants' rumination traits revealed that the colouring scheme had a significant effect on the rumination score, $F_{1,157} = 9.53$, $p < .01$, Bonferroni corrected. Post-hoc analysis revealed that Multicoloured (MC) visualisations lead to higher scores on the rumination scale than single-coloured (SC) visualisations, $p < .05$. No effect of shape used was found, nor were any interaction effects present. Figure 4 depicts the results showing rumination score given colouring scheme (MC, SC).

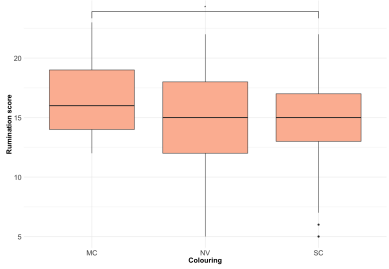


**Figure 4: Rumination scores for the colour schemes.**

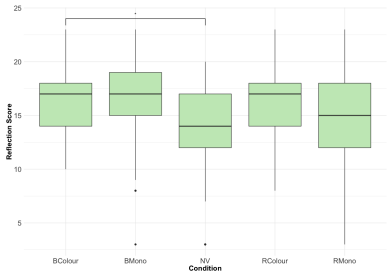


**Figure 5: Reflection scores for the five experimental conditions.**

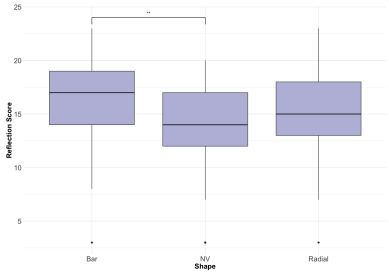


**Figure 6: Reflection scores given the shapes used.**

## 7.2 Reflection

Accordingly, we conducted a one-way ANCOVA to investigate the effect of the type of visualisation used on the reflection scale score, controlling for the participants' reflection traits. We found a significant effect, $F_{4,158} = 2.94$, $p < .05$. Post-hoc testing with Tukey's HSD revealed significant differences between the condition pairs NV — BColour, with $p < .05$. Figure 5 shows the results. A significant effect of the covariate, reflection trait scale score was present, $F_{1,158} = 7.54$, $p < .01$. As in the case of rumination, we investigated the effect of the colouring scheme (multicoloured, single-coloured) and shape used (radial bar chart, bar chart) on the reflection score. A two-way ANCOVA, controlling for the participants' reflection trait revealed that the shape used had a significant effect on the reflection score, $F_{2,158} = 5.40$, $p < .01$, Bonferroni corrected. Tukey's HSD revealed that bar charts lead to higher scores on the reflection scale at $p < .05$. No effect of colour scheme was found, nor were any interaction effects present. Figure 6 depicts the results showing reflection score given shape used.

| P | Gender | Age | Visualisation Prototype | Education |
|---|---|---|---|---|
| P1 | Female | 24 | BMono | Master |
| P2 | Female | 21 | RMono | Bachelor |
| P3 | Male | 23 | BColour | High School |
| P4 | Male | 24 | RColour | Master |
| P5 | Female | 23 | RMono | Bachelor |
| P6 | Male | 22 | BColour | High School |
| P7 | Female | 22 | RColour | High School |
| P8 | Female | 20 | BMono | High School |
| P9 | Female | 49 | RMono | Bachelor |
| P10 | Male | 51 | BMono | Master |
| P11 | Male | 23 | BColour | Bachelor |
| P12 | Male | 23 | RColour | Bachelor |
| P13 | Male | 22 | BMono | High School |
| P14 | Female | 22 | RMono | Bachelor |
| P15 | Female | 23 | BColour | High School |
| P16 | Male | 34 | RColour | Master |
| P17 | Female | 21 | BMono | High School |
| P18 | Female | 37 | RMono | Master |
| P19 | Female | 22 | BColour | High School |
| P20 | Female | 26 | RColour | Master |

**Table 3: Details of participants P in our interview study.**

# 8 STUDY 2: INTERVIEWS

Having observed that the way of presenting unmet goals could affect reflection and rumination, we wanted to understand the reasons behind this effect. To that end, we decided to conduct semi-structured interviews to explore the users' thoughts when confronted with the visualisations. Each participant was randomly assigned one of the four visualisations used in Study 1, which was the main topic of the interview. The interviews lasted an average of 25.68 minutes ($MIN = 17$–$MAX = 35$). Each of the interviews was a one-on-one session with a single researcher.

## 8.1 Participants

We recruited $n = 20$ participants, aged 20–51 ($M = 26.6$, $SD = 9.03$), with 8 males and 12 females. The participants were asked for consent for recording before the interviews and informed that they are allowed to terminate the interview at any time. All participants confirmed that they did not have red-green colour blindness. The participants were recruited through social media posts, flyers at the university canteen and snowball sampling. We provided an online shopping voucher worth $20 as compensation. Table 3 presents details about the participants.

## 8.2 Interview Protocol

In the interview, we first obtained demographic data. We then presented one of the four visualisations (BMono, BColour, RMono, RColour) to the participants and asked them to imagine that the prototypes presented their previous week's activity and think aloud while interpreting the data [4]. Based on what the participants voiced during the think-aloud protocol we then inquired in more detail about aspects such as form, colour and emotions connected to the visualisation. In the final part of the interview, we gave the participants the opportunity to ask follow-up questions, thanked them for their participation in our study and provided them compensation information.

## 8.3 Analysis

All audio recordings were transcribed verbatim and imported into the Atlas.ti analysis software. We applied open coding combined with thematic analysis, as described by Blandford et al. [8]. Two researchers coded a representative sample of 15% of the material. Next, a coding tree was established through iterative discussion. The remaining transcripts were split between the two researchers and coded individually. A final discussion session was conducted to structure the coding tree after the material was coded. We then identified emerging themes from the material: Form and Colour, Reflection and Rumination and Self-compassion and Performance Pressure.

# 9 STUDY 2: RESULTS

Here we provide a detailed description of the themes and illustrate them with excerpts from the interview data. Our qualitative results challenge the binary view regarding fitness tracker feedback, especially regarding colour in the fitness tracker visualisation prototypes and goal setting.

We found that participants were able to relate to the presented prototypes. Many participants provided unsolicited comments, reflecting on similarities and differences between the presented visualisation and the visualisations they experienced in their past technology use. Furthermore, interviewees contemplated about their own days and how they could be mapped to the presented data. This was illustrated by the following quote:

*On Sundays I'm always in the animal shelter, but there I also get a lot of steps. (...) This (the visualisation) is amazingly close to my weekly routine. It really looks like my week. (P19, BColour)*

## 9.1 Form and Colour

Bar charts were perceived as intuitive by all participants presented with this visualisation. Many interviewees explained the visualisation without hesitation. One participant remarked that it was easy to understand and aesthetically pleasing:

*It's very minimalist, but I find it easy to understand. I know exactly what it tells me. (...) So there are no axes, which I find actually more beautiful, because there is just less on the screen. I really like the minimalist things. You have an obvious line, so I find it very understandable. (P10, BMono)*

In contrast, the group of participants presented with the radial bar chart often expressed confusion. The participants suggested that this visualisation may have a steeper learning curve:

*Well, if you know it, you'll probably get used to it, but you have to know where to look first. (...) What they (the designers) did not do well is that the rings are hard to map to values. Yes, so somehow I find it really confusing. (P18, RMono)*

Some of the participants primarily commented on the colours and disregarded the form of the visualisation. For many participants, the colours were perceived as self-explanatory and red was associated

with negative feelings. Some participants even perceived the colour red as depressing:

*So the colours are relatively logical. Green is mostly positive, red is mostly negative (...), you just want to see more green. And red is a bit depressing. (...) I do not want to constantly look at my goals and always be reminded of what I did not manage, because everything is red. (P7, RColour)*

The emotional and motivational consequences of the colour scheme were discussed by many participants. Interestingly, some interviewees reflected on the negative perception of the colour red, but, at the same time, mentioned that the highlighted failure would motivate them:

*Red is bad. It's just a harsh colour, and it's also the colour that describes bad or below-average days. The colour appears aggressive. Perhaps the designers considered that, perhaps they wanted to express caution with the red colour. (...) For me, it would have a motivating effect, if the data looked bad. But I guess that depends on the personality. (P12, RColour)*

## 9.2 Reflection and Rumination

Another recurrent theme in the collected data described reflection and rumination processes which go beyond reflection and rumination triggered by the specific visualisation prototypes. For instance, one participant emphasised their personal interest in contrasting their personal perception of the physical activity level with the numbers presented by the fitness tracker:

*(...) it is always interesting to see how much you have really moved versus your perception. And I always find it very interesting, because somehow 11 000 steps does not sound like that much, if you only see the number. But when you think, "Okay, that's how much time or kilometres that is". (P5, RMono)*

For some participants, the idea of wearing a fitness tracker induced too much pressure and made them uncomfortable. One participant even stopped tracking, because the tracker did not reflect the actual fitness level, which lead to stressful situations:

*(...) It has put me under some pressure. For example, when I was cycling as a sports activity, the pedometer did not register it. That said, I had very few steps at the end of the day, even though my physical activity was actually very high, and that put me under pressure at some point. (P2, RMono)*

The ability to react to negative feedback of the fitness tracker was a recurrent topic. The time that was left to counterbalance negative feedback was perceived as crucial and if there was not enough time left to work on reaching the goal, it had negative emotional consequences for the user:

*It depends on how much time I have left. If I have plenty of time left to balance the reds, then it's such an inciting, stimulating thing. When I realise, "Oh dear, now I do not have enough time left to do this", it makes me grumpy. (P3, BColour)*

Experiencing negative emotional consequences because of goal failure is also highlighted by this statement:

*(...) If the tracker shows me "Okay, you only walked 3000 steps today"; then I always have a very bad conscience, so that's the incentive. (P20, RColour)*

## 9.3 Self-Compassion and Performance Pressure

This theme describes two contrasting notions. Here, participants discussed the need to do something that goes against the recommendations from their fitness tracker, such as taking a break. Some participants also discussed how the tracker did not reflect their day properly or commented about performance pressure and the need to reach their goal. Interestingly, they emphasised that it was particularly important to them to reach their goal when they were very close to it:

*So, in the evening, when I see: 'Okay, I am now at, what do I know, 10700' then I walk around the block one more time thinking: 'Okay, now you are so close to the 11 000, you can do that now.' But when I see it is completely far away, then I do not really care about it. (P34, BColour)*

While most participants described that it would feel uncomfortable to not reach their goal, some interviewees made a point of being compassionate towards themselves:

*(...) because as I said before, I'm quite active overall and then it is not bad if one day it is not there. Although this looks like a warning now, but you do not have to be so active every day. That's fine too. (P4, RColour)*

The struggle between self-compassion and performance pressure is illustrated by the following comment, where an interviewee reflected about the question of whether it was okay to relax or if that would be lazy:

*I could say: 'Oh no, I have to do something'. But maybe I do not have to do anything. But then that leads to a certain tension, where one says 'Do I have to do more? Do I need that?'. Is it laziness when you say you do not need that? Because there are already situations where you say you need relaxation. But then you do not do that too often. I think there's a bit of pressure or motivation to maintain your performance. (P18, RMono)*

# 10 DISCUSSION

In this research, we explored possible reflection and rumination triggered by four different visualisations and a textual description of an unmet fitness tracker goal. We conducted two studies, an experimental vignette study and an interview study. Our results challenge the generally positive view of providing goal feedback. In our experimental vignette study, we observed that multicoloured visualisations led to significantly more rumination (**RQ1**). While the affective consequences of the multicoloured charts were mainly described as negative in our interviews, some interviewees perceived them as uncomfortable but motivating at the same time, cf. Form and Colour. Our results highlight the importance of fitness tracker design to be aware of potential negative consequences of multicoloured graphs and the possibility of users ruminating over goal failure. More precisely, our qualitative data shows that the users exhibit a variety of attitudes in how they relate to the presented goal achievement data. That is why current fitness trackers should present alternative goal visualisation options to users so that the users can choose an interface matching their goal orientation. Allowing the user to make their own choice potentially satisfies

the need for autonomy. Further, our results suggest that a neutrally coloured visualisation should be the default with coloured options being available after informing the user of possible negative consequences.

A tangible recommendation that stems from our results is that *fitness trackers should be wary about using strongly pre-attentive attributes like colour schemes to highlight failure metrics, such as unmet goals. Trackers should make sure that such intense feedback does not lead to negative affective consequences.*

In line with approaches known from Positive Psychology, fitness tracker visualisations could instead focus on goal attainment and de-emphasise goal failure. This, in turn, could foster a 'growth mindset', i.e. empowering people to build on their personal resources and increase resilience to cope with eventual failure, such as not achieving a goal [24]. This resonates with the findings from our interviews were some participants articulated feeling stressed by negative feedback. Improving the resilience in dealing with challenging situations through fostering a 'growth-mindset' could potentially support fitness tracker users in enjoying their success when achieving a goal. Furthermore, the induced positive feelings and the increased resilience could help users to utilise the advantages of goal setting mechanisms in current fitness apps, while minimising the negative effects of them, such as negative thought cycles or feeling pressured.

We found that multicoloured graphs and bar graphs triggered significantly more reflection (**RQ2**) in our quantitative inquiry. Our interview results showed that the multicoloured charts triggered more rumination than the single-coloured visualisations. One possible explanation for this effect is illustrated by our qualitative results. Participants commented extensively about unmet goals highlighted through orange and red colouring in the Form and Colour theme. They also highlighted situations were the negativity of red was unwarranted in the Self-Compassion and Performance Pressure theme. This is an intriguing finding, since it showcases the complexity of designing rich fitness tracker visualisations that can potentially foster reflection but, at the same time, are not judgemental, i.e. they reduce pressure and potential rumination in their users. Consequently, future fitness tracker should balance how feedback with respect to goal is communicated, allowing user to customise how goals are displayed, depending on their goal orientation.

While we observed some significant differences in the study, our work also shows that it is likely that the design space of visualisations that avoid the negative effects of tracking is considerably large. This implies that designers of fitness tracking applications have a wide variety of choices for the visualisations to be used for a particular product. This creates an opportunity for development beyond the current static, pre-determined visualisation designs for personal informatics data views (e.g. the Fitbit weekly report, which uses radial bar charts for goals). As different users exhibit different levels of reflection and rumination when presented with different visualisations, users should be empowered to make their own visualisation choices. Consequently, future fitness tracker applications should offer alternative visualisations that the users can interactively change.

Based on our results and given that users are more familiar with bar charts than with radial bar charts, we hypothesise that the bar charts led to more positive affect. This assumption was confirmed by our interview results in the Form and Colour theme. Participants commented extensively on both chart types (radial bar chart and bar chart) and all participants perceived the bar chart as very intuitive. In contrast, some participants reflected on the aesthetically pleasing qualities of the radial bar chart but at the same time described it as confusing and not intuitive.

## 10.1 Ways Forward

Based on our results, we highlight possible ways forward for designing fitness tracker experiences. Our work shows that the design properties of the visualisations used for fitness tracker data may affect the way users connect to goals and translate their qualitative goals to tracker data. In line with the *Tracker Goal Evolution Model* [43], one way to address the challenge of evolving and occasionally unmet fitness tracker goals is to shift the focus of users towards what they have already achieved; more precisely, to foster a growth mindset [18]. This resonates with past work which stressed the importance of enabling users to connect their fitness tracker goals to their higher level goals [30, 43]. Our studies focused on the visualisation of numeric goals. Future work can explore how users can be supported to connect quantitative and qualitative goals and how that might help to mitigate rumination.

Also, current fitness trackers collect a variety of different data, such as steps, heart rate and sleep quality. Future research can explore if our findings can be replicated when exploring visualisations of fitness tracker data beyond steps, such as sleep quality and heart rate. Yet, other metrics come with more complexity, e.g. meeting sleep quality goals is harder to assess than determining if a specific number of steps was taken. Hence, a question for future research could be: How to foster reflection and avoid rumination when goal achievement is hard to define. Future work could also study how other feedback modalities such as temporal aspects, audio or haptic feedback might influence rumination and reflection of users.

The need for tracker visualisations to reflect daily activities and life circumstances is emphasised by the findings in our interviews, especially under Self-Compassion and Performance Pressure. Some participants anticipated that they would struggle to find a balance between reaching their fitness tracker goals, following the feedback provided their device and doing what felt healthy and meaningful to them (e.g. working on one's thesis or taking a break). Interviewees reflected that their fitness tracker may have reduced complex activities to numbers, which did not reflect if they had had an active, productive day. Hence, incorporating self-compassion into fitness tracker design has the potential to enable fitness tracker users to find a balance between their higher level fitness goals and their higher level wellbeing goals, without the need for the tracker to process detailed data about the nature of the user's activities.

Even though we hope that our work can support the design of future fitness trackers that fosters reflection and avoids triggering rumination, we are convinced that studying how to communicate failure is only one way to address this topic. We believe that future research should not only focus on how best to communicate failure, but explore why 'failure' occurs in the first place. Current personal informatics systems are mainly designed for and with the white, healthy, mentally stable user in mind [51]. For instance, users who already struggle with physical or mental challenges are already

prone to potentially ruminate more [50]. Thus, in line with Spiel et al. [51], we emphasise that building more inclusive personal informatics systems is key for engaging, meaningful, long-term tracking experiences. Consequently, future systems should develop customisation strategies that adapt visualisations to users, preventing failure and reducing negative thought cycles.

### 10.2 Limitations

Our work constitutes a first step towards exploring ways how to communicate unmet fitness tracker goals through visualisations. Yet, we recognise that the approach used in this paper is prone to certain limitations. It has to be noted that the hues of the multicoloured bars may not be discernible by people who have a colour vision deficiency. It was not possible for us to check reliably if the participants in the experimental vignette study have a colour vision deficiency. However, we only included participants with no colour vision deficiency in our interview study.

Furthermore, we used hypothetical data to study the potential effects of different ways to communicate failure. We opted for this approach for ethical reasons. Previous work showed that results derived with vignette studies generate similar results than real experiences (e.g, [21, 31]). Furthermore, within our interviews, participants related the prototypes to real experiences and contextualised what they saw with socially relevant experiences of their own lives. This approach represents a trade off between ecological validity and what is ethically possible. Consequently, even though we believe that our approach was justified as we were striving for an ethically sound approach and high internal validity, future research should explore if our findings can be replicated when studying the visualisations with the real fitness tracker data of participants. For instance, future research could utilise longitudinal, qualitative studies to explore the reasons behind participants' preferences for certain visualisations over a longer period of time.

Moreover, we can not exclude the possibility that other chart designs than those chosen for our study would have led to other results. However, as this is (to the best of our knowledge) the first exploration of this issue, we aimed to evaluate commercially available chart types, for our work to be relevant to designers and users relating to existing choices. The two types, bar charts (commonly considered to be one of the most basic and understandable charts) and radial bar charts (less well known and less easily parsed), offered a good opportunity to test the effects of chart type. Past work in visualisation for personal informatics [33] pointed to differences between traditional and more elaborate visualisations and our work explores this question. Any experimental design is a balance between variety (number of choices to test) and reliability (number of participants per choice), and we chose higher reliability.

## 11 CONCLUSION

This paper investigated the effects of different ways to communicate unmet fitness tracker goals through visualisation on rumination and reflection. We conducted two studies; a between-subject online study with five conditions — four different prototype visualisations and a textual description of unmet fitness tracker goals and a qualitative interview study. We found that bar graphs offered a significantly better potential for reflection and multicoloured charts triggered significantly more rumination. We further investigated the differences between the prototypes in semi-structured interviews. The interviews showed that users were concerned about how the visualisations reflected their activity. We also observed that the colour red produced emotionally-charged responses which could both provide motivation and lead to rumination.

Our results imply that future trackers should make sure that very direct feedback of unmet goals does not lead to negative affective consequences, such as rumination, for the user. Potential ways forward to address this challenge could be to strictly using single-coloured visualisations as the default data view, fostering a growth-mindset in the participants and supporting self-compassion through fitness tracker design. We provided insights into the design of future fitness trackers that account for occasionally unmet fitness tracker goals and foster a meaningful long-term experience. We hope that our results will inspire further research into alternative designs for personal informatics visualisations that ensure that negative effects of tracking are minimised.

## REFERENCES


[1] Henk Aarts, Ruud Custers, and Rob W Holland. 2007. The nonconscious cessation of goal pursuit: When goals and negative affect are coactivated. *Journal of personality and social psychology* 92, 2 (2007), 165.

[2] 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.

[3] Herman Aguinis and Kyle J Bradley. 2014. Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods* 17, 4 (2014), 351–371.

[4] Obead Alhadreti and Pam Mayhew. 2018. Rethinking Thinking Aloud: A Comparison of Three Think-Aloud Protocols. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 44, 12 pages. https://doi.org/10.1145/3173574.3173618

[5] World Medical Association et al. 2013. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama* 310, 20 (2013), 2191.

[6] 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

[7] Eric PS 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*. ACM, 93–102.

[8] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. *Qualitative HCI research: Going behind the scenes*. Vol. 9. Morgan & Claypool Publishers. 1–115 pages.

[9] T. Blascheck, L. Besançon, A. Bezerianos, B. Lee, and P. Isenberg. 2019. Glanceable Visualization: Studies of Data Comparison Performance on Smartwatches. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (Jan 2019), 630–640. https://doi.org/10.1109/TVCG.2018.2865142

[10] Patricia C Broderick. 1998. Early adolescent gender differences in the use of ruminative and distracting coping strategies. *The Journal of Early Adolescence* 18, 2 (1998), 173–191.

[11] Kirk Warren Brown and Richard M Ryan. 2003. The benefits of being present: mindfulness and its role in psychological well-being. *Journal of personality and social psychology* 84, 4 (2003), 822.

[12] James Clawson, Jessica A. Pater, Andrew D. Miller, Elizabeth D. Mynatt, and Lena Mamykina. 2015. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. ACM Press, 647–658. https://doi.org/10.1145/2750858.2807554

[13] William S Cleveland and Robert McGill. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American statistical association* 79, 387 (1984), 531–554.

[14] Sunny Consolvo, Predrag Klasnja, David W McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A Landay. 2008. Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 54–63.

[15] Sunny Consolvo, Predrag Klasnja, David W McDonald, and James A Landay. 2009. Goal-setting considerations for persuasive technologies that encourage physical activity. In *Proceedings of the 4th international Conference on Persuasive Technology*. ACM, 8.

[16] Tim Coughlan, Michael Brown, Glyn Lawson, Richard Mortier, Robert J Houghton, and Murray Goulden. 2013. Tailored scenarios: a low-cost online method to elicit perceptions on designs using real relationships. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 343–348.

[17] Frederick E. Croxton and Harold Stein. 1932. Graphic Comparisons by Bars, Squares, Circles, and Cubes. *J. Amer. Statist. Assoc.* 27, 177 (1932), 54–60. https://doi.org/10.1080/01621459.1932.10503227 arXiv:https://amstat.tandfonline.com/doi/pdf/10.1080/01621459.1932.10503227

[18] Carol S Dweck. 2006. Mindset: The new psychology of success." New York: RandomHouse. (2006).

[19] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, 667–676.

[20] 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.

[21] Spencer C Evans, Michael C Roberts, Jared W Keeley, Jennifer B Blossom, Christina M Amaro, Andrea M Garcia, Cathleen Odar Stough, Kimberly S Canter, Rebeca Robles, and Geoffrey M Reed. 2015. Vignette methodologies for studying clinicians' decision-making: validity, utility, and application in ICD-11 field studies. *International journal of clinical and health psychology* 15, 2 (2015), 160–170.

[22] Stephen Few. 2006. Information dashboard design. (2006).

[23] Ayelet Fishbach, Tal Eyal, and Stacey R Finkelstein. 2010. How positive and negative feedback motivate goal pursuit. *Social and Personality Psychology Compass* 4, 8 (2010), 517–530.

[24] Barbara L Fredrickson. 2001. The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American psychologist* 56, 3 (2001), 218.

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

[26] Mitchell Gordon, Tim Althoff, and Jure Leskovec. 2019. Goal-setting and achievement in activity tracking apps: A case study Of MyFitnessPal. In *The World Wide Web Conference*. 571–582.

[27] R Gouveia, E Karapanos, and M Hassenzahl. 2015. Activity trackers in the wild: Design strategies and their impact on engagement and physical activity in daily life. *M-ITI Technical Report* (2015).

[28] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A Munson, and Marc Hassenzahl. 2016. Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 144–155.

[29] James J Gross and Oliver P John. 2003. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology* 85, 2 (2003), 348.

[30] 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

[31] Jens Hainmueller, Dominik Hangartner, and Teppei Yamamoto. 2015. Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences* 112, 8 (2015), 2395–2400.

[32] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 203–212.

[33] 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 (2015), 420–433.

[34] Dmitri S. Katz, Blaine A. Price, Simon Holland, and Nicholas Sheep Dalton. 2018. Data, Data Everywhere, and Still Too Hard to Link: Insights from User Interactions with Diabetes Apps. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, Article 503, 12 pages. https://doi.org/10.1145/3173574.3174077

[35] Rohit Ashok Khot, Deepti Aggarwal, Ryan Pennings, Larissa Hjorth, and Florian'Floyd' Mueller. 2017. Edipulse: investigating a playful approach to self-monitoring through 3D printed chocolate treats. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 6593–6607.

[36] 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.

[37] Joyca Lacroix, Privender Saini, and Roger Holmes. 2008. The relationship between goal difficulty and performance in the context of a physical activity intervention program. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. ACM, 415–418.

[38] 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.

[39] Edwin A Locke and Gary P Latham. 1985. The application of goal setting to sports. *Journal of Sport and Exercise Psychology* 7, 3 (1985), 205–222.

[40] Winter Mason and Siddharth Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods* 44, 1 (2012), 1–23.

[41] Alain Morin. 2017. Toward a glossary of self-related terms. *Frontiers in psychology* 8 (2017), 280.

[42] Alistair Morrison and Viktor Bakayov. 2017. Stickers for steps: a study of an activity tracking system with face-to-face social engagement. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017).

[43] 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.

[44] Judith S Olson and Wendy A Kellogg. 2014. *Ways of Knowing in HCI*. Vol. 2. Springer.

[45] Gabriele Paolacci, Jesse Chandler, and Panagiotis G Ipeirotis. 2010. Running experiments on amazon mechanical turk. (2010).

[46] W. Playfair and James Corry. 1786. *The Commercial and Political Atlas: Representing, by Means of Stained Copper-plates, the Exports, Imports, and General Trade of England; the National Debt, and Other Public Accounts; with Observations and Remarks*. J. Debrett. https://books.google.se/books?id=aav9tAEACAAJ

[47] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers Chalmers. 2014. Personal tracking as lived informatics. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 1163–1172.

[48] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[49] Jeannette M Smith and Lauren B Alloy. 2009. A roadmap to rumination: A review of the definition, assessment, and conceptualization of this multifaceted construct. *Clinical psychology review* 29, 2 (2009), 116–128.

[50] Heather Soo, Sue Burney, and Christopher Basten. 2009. The role of rumination in affective distress in people with a chronic physical illness: A review of the literature and theoretical formulation. *Journal of Health Psychology* 14, 7 (2009), 956–966.

[51] Katta Spiel, Fares Kayali, Louise Horvath, Michael Penkler, Sabine Harrer, Miguel Sicart, and Jessica Hammer. 2018. Fitter, Happier, More Productive?: The Normative Ontology of Fitness Trackers. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. ACM, New York, NY, USA, Article alt08, 10 pages. https://doi.org/10.1145/3170427.3188401

[52] 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

[53] Dorthe Kirkegaard Thomsen, Jan Tønnesvang, Anette Schnieber, and Martin Hammershøj Olesen. 2011. Do people ruminate because they haven't digested their goals? The relations of rumination and reflection to goal internalization and ambivalence. *Motivation and Emotion* 35, 2 (2011), 105–117.

[54] Paul D Trapnell and Jennifer D Campbell. 1999. Private self-consciousness and the five-factor model of personality: distinguishing rumination from reflection. *Journal of personality and social psychology* 76, 2 (1999), 284.

[55] Catrine Tudor-Locke and David R Bassett. 2004. How many steps/day are enough? *Sports medicine* 34, 1 (2004), 1–8.

[56] Shaun Wallace, Hua Guo, and David Sasson. 2017. Visualizing Self-Tracked Mobile Sensor and Self-Reflection Data to Help Sleep Clinicians Infer Patterns. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. ACM, New York, NY, USA, 2194–2200. https://doi.org/10.1145/3027063.3053138

[57] Colin Ware. 2012. *Information visualization: perception for design*. Elsevier.

[58] Edward R Watkins. 2008. Constructive and unconstructive repetitive thought. *Psychological bulletin* 134, 2 (2008), 163.

[59] Paweł Woźniak, Kristina Knaving, Staffan Björk, and Morten Fjeld. 2015. RUFUS: Remote Supporter Feedback for Long-Distance Runners. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 115–124. https://doi.org/10.1145/2785830.2785893