Check for updates

Designing Data Visualisations for Self-Compassion in Personal Informatics

MEAGAN B. LOERAKKER, Chalmers University of Technology, Sweden
JASMIN NIESS, University of Oslo, Norway
MARIT BENTVELZEN, Utrecht University, the Netherlands
PAWEŁ W. WOŹNIAK, Chalmers University of Technology, Sweden

Wearable personal trackers offer exciting opportunities to contribute to one's well-being, but they also can foster negative experiences. It remains a challenge to understand how we can design personal informatics experiences that help users frame their data in a positive manner and foster self-compassion. To explore this, we conducted a study where we compared different visualisations for user-generated screen time data. We examined positive, neutral and negative framings of the data and whether or not a point of reference was provided in a visualisation. The results show that framing techniques have a significant effect on reflection, rumination and self-compassion. We contribute insights into what design features of data representations can support positive experiences in personal informatics.

CCS Concepts: • **Human-centered computing → HCI theory, concepts and models**.

Additional Key Words and Phrases: Personal informatics, self-compassion, reflection, rumination, data visualisation

ACM Reference Format:
Meagan B. Loerakker, Jasmin Niess, Marit Bentvelzen, and Paweł W. Woźniak. 2023. Designing Data Visualisations for Self-Compassion in Personal Informatics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 169 (December 2023), 22 pages. https://doi.org/10.1145/3631448

1 INTRODUCTION

Most fitness tracker experiences on the market are designed with goal completion at the centre [30]. This, in turn, leads to tracking being focused on performance as ambitious goal setting promotes high achievement [39]. Yet, goal-focused tracking is prone to a number of pitfalls. A person's confidence in their ability to succeed in their goals can be undermined as a result of negative feedback [61]. Users may abandon goals [1] altogether if dissatisfied. Experiences of not fulfilling goals can negatively affect well-being and trigger negative thought cycles. Niess et al. [53] found that visualisations highlighting the failure to achieve a goal compared to more neutral visualisations triggered negative affect in fitness tracker users. Consequently, it remains a challenge for Human-Computer Interaction (HCI) to understand how negative thought cycles can be prevented through designing personal tracking tools. In order for tracking systems to continue to contribute to users' well-being, we need to explore the design elements in personal tracking applications which allow for communicating goal failure and completion in ways which positively affect the user.

Previous work in HCI shows that it remains a challenge to create meaningful and supportive experiences in Personal Informatics (PI) [10, 31]. In response to this, past research focused on designing personal informatics systems that enhance self-reflection (e.g. [3, 15, 19, 28, 32, 34, 37, 38, 43, 44, 63, 66]), as this is seen as a key enabler

Authors' addresses: Meagan B. Loerakker, meagan@chalmers.se, Chalmers University of Technology, Gothenburg, Sweden; Jasmin Niess, jasminni@uio.no, University of Oslo, Oslo, Norway; Marit Bentvelzen, m.bentvelzen@uu.nl, Utrecht University, Utrecht, the Netherlands; Paweł W. Woźniak, pawel.wozniak@chalmers.se, Chalmers University of Technology, Gothenburg, Sweden.

2474-9567/2023/12-ART169
https://doi.org/10.1145/3631448

for positive experiences in a user's personal informatics journey [10]. However, as noted by Eikey et al. [22], when reflective thought occurs, it sometimes leads to negative thought and emotion cycles, also known as rumination. While previous work has indicated that the way that personal data is represented can trigger rumination or stress (e.g. [37, 53]), it remains an open question how to design PI interfaces that prevent rumination.

Here, we explore one way to reduce rumination in PI—self-compassion. In 2020, Niess et al. [53] proposed self-compassion as a potential approach to assist users of PI systems in reflecting on their personal data, while also reducing rumination. In 2021, this was also echoed by Eikey et al. [22]. Along similar lines, they noted that targeting users' emotional experiences could offer a possibility for reducing rumination, which could be accomplished through increasing self-compassion. Self-compassion involves 'self-directed' care and an understanding that we are all human, struggle and have to cope with challenging emotions (e.g. [48, 56]). Studies by Neff [47, 52] demonstrated that higher self-compassion levels are associated with lower levels of rumination. In line with this, Samaie and Farahani [62] found an association between the psychological concepts of rumination and self-reflection: self-compassion amplifies the relation between self-reflection and stress, but it also impairs the relationship between rumination and stress. In other words, self-compassion can play a moderating role in the relationship between stress, and the cognitive processes rumination and self-reflection [62]. Samaie and Farahani's [62] study showed that a higher level of rumination is associated with more stress, and that higher levels of self-reflection are associated with less stress. This implies that the higher the self-compassion levels, the more the individual's self-reflection protects against stress, and the more the link between an individual's rumination and stress is disrupted [62]. Furthermore, self-compassion may reduce the experience of negative emotions. For example, Leary et al. [40] suggest that those who are more self-compassionate may experience fewer negative emotions. Leary et al. [40] suggested that those who are more self-compassionate are less likely predisposed to self-ruminate in the aftermath of a negative event.

These findings suggest that designing for self-compassion opens up opportunities for both encouraging positive cognitive processes and discouraging negative ones (e.g. [40, 47, 52, 62]), due to its alleged moderating effect. However, it remains unclear how we can design PI systems that moderate rumination by enhancing self-compassion. To address this gap, we explored how different personal data representations influenced participants' levels of reflection, rumination and self-compassion. We use framing techniques as a starting point to design for self-compassion, which is inspired by the well-known 'Framing Effects' defined by Tversky and Kahneman [68]. Framing has been shown to successfully foster positive feelings and cognitive processes like self-efficacy [18], or positive behaviours like avoiding to install an app with questionable privacy policies [17]. Despite these promising results, there is still a need to explore if and how framing can support self-compassionate mindsets. To that end, we conduct a within-subject study with $n = 18$ participants of how users explore different visualisations (with either negative, neutral or positive framing) of their screen time data.

This paper contributes the following: (1) results on how personal data visualisations influence users' reflection, rumination, and self-compassion levels, and (2) insights on how the design on personal informatics experiences can foster self-compassion. In this work, we first review related work on self-compassion, self-reflection, and rumination. We relate these findings to previous work on PI systems to inform the design of the data visualisations used in the study. We then present the method and findings from our study, and conclude with a discussion and directions for future research.

2 RELATED WORK

Here, we first review related work on personal informatics (PI) with a focus on reflection and rumination. We then discuss the concept of self-compassion and contextualise it within the field of PI.

2.1 Self-Reflection and Rumination in Personal Informatics

Self-reflection is considered to be a central part of personal informatics systems. Such systems enable users to understand their health and well-being through automatically collected personal data. By reflecting on personal data, a user can notice patterns and trends, which can lead to more knowledge about oneself [25].

Several models have been proposed to describe the user's journey in using personal informatics systems. These models aim to provide a more detailed understanding of the reflection process. One such model is the *Lived Informatics Model of Personal Informatics* proposed by Epstein et al. [27], which is an extension of Li et al.'s [42] *Stage-Based Model of Personal Informatics Systems*. The Lived Informatics Model consists of four stages: deciding to track, selecting tools, tracking and acting (which is an ongoing process of collection, integration, and reflection), and lapsing. The Lived Informatics Model of Personal Informatics was later extended by Bentvelzen et al. [10]. They introduced the *Technology-Mediated Reflection Model* (TMRM) [10] to describe user behaviours and practices during the reflection phase of a personal informatics journey. The model shows how users enter, exit and stay in the reflection phase. In the TMRM, reflection is considered to be a temporary state—a dynamic process in which a user constantly adapts their tracking experience to their evolving needs. It has been argued that fitness tracker users struggle to connect qualitative life goals with the numeric tracker goals, as shown in Niess and Woźniak's [54] *Tracker Goal Evolution Model*. In their model, reflection is described as a means to translate those qualitative goals into quantitative ones [54], whereas other models describe reflection as a part of the tracking process (e.g. the Lived Informatics Model [27] and the Stage-Based Model [42]). Interestingly, the concept of self-compassion is not featured in personal informatics models thus far.

Recently, scholars in HCI have started engaging with the complexities of reflection. For instance, a paper by Niess et al. [53] highlights a potential risk for reflection-enhancing systems. They postulate that users of such systems can potentially experience negative thought cycles. The authors introduce the psychological term *rumination* to refer to these cycles. Their empirical work showcased that rumination can be experienced by users of personal informatics systems. This was later echoed by Eikey et al. [22]. Rumination and reflection are related concepts [67] as both concepts refer to the process of the person focusing their attention upon themselves [46]. The main difference between the two concepts, is the fact that reflection is usually regarded as a positive process [67] as it revolves around curiosity [53], whilst rumination has a negative connotation to it as it revolves around the focus on one's loss or failures [46, 64]. While rumination can undermine technology for reflection, it remains an open question of how to balance promoting reflection on the one hand and preventing rumination on the other. According to Eikey et al. [22], targeting users' emotional experiences could offer a possibility for reducing rumination, which could be accomplished through increasing self-compassion. Therefore, in this work, we explore how personal informatics systems can foster self-compassion and how this affects reflection and rumination.

According to the conceptual work by Eikey et al. [22], emotion is one of the three mechanisms that explain how rumination and PI tools are linked. With this link, they explain how designing for emotion is one of the ways to reduce rumination [22]. Increasing self-compassion was mentioned as one of the methods for improving the emotional experience with PI tools [22, 53]. Since it has been shown that self-compassion is positively associated with life satisfaction and negatively associated with rumination and other negative processes and feelings [47, 51], designing for self-compassion in PI tools shows potential to prevent rumination in users. However, more research is needed in order to gain a shared understanding as to how we can balance the promotion of self-reflection and the mitigation of rumination [53] and this works explored possible design elements for mitigating rumination.

2.2 Self-Compassion

In previous work, the concept of self-compassion is described as 'an especially adaptive self-attitude due to its positive association with multiple aspects of psychological well-being' [16, p. 1133]. Self-compassion is a cognitive process that often gets connected with Buddhism (e.g. [56]). Over the years, Eastern constructs, such as Buddhism, have gained traction in therapeutic developments like *mindfulness-based cognitive therapy*, *acceptance and commitment therapy*, and *compassion-focused therapy* (CFT) due to their positive outcomes [5, 29, 33, 36]. Kristin Neff, one of the pioneers in the study of self-compassion, derived her definition of self-compassion from Buddhism as well [49]. According to Neff [48], self-compassion consists of three components: (1) self-kindness, (2) common humanity, and (3) mindfulness. Neff [48] defines self-kindness as the extension of kindness, rather than harsh self-criticism, aimed towards oneself [48]. They define common humanity as the action of putting one's experiences into perspective of the human experience, and according to Neff [48] mindfulness is the idea of balancing negative thoughts and emotions [48]. Self-compassion introduced a new research perspective to studying optimal psychological functioning, as it presents a more concrete integration of the concern for oneself and others [14]. Compassion towards others involves offering them 'patience, kindness and non-judgmental understanding, recognising that all humans are imperfect and make mistakes' [47, p. 224], whereas compassion towards oneself involves recognising that one's inadequacies and experiences are part of the human experience [48]. Although these definitions of self-compassion are grounded in theory in the field of Psychology, more research is needed to gain a deeper understanding of the concept's individual components [6].

These suggestions and findings on self-compassion, stress and rumination support Samaie and Farahani's [62] idea that rumination might be a moderating factor in the link between self-compassion and stress. In line with this, Niess et al. [53] found that rich data visualisations, as used in fitness trackers, can both foster reflection and reduce potential rumination by making the visualisations not as judgemental. This suggests an opportunity for PI tools to incorporate self-compassion in order to help users find their balance between their fitness and well-being goals [53]. Designing for self-compassion provides an opportunity for more positive experiences with PI tools and the exploration of personal data. Furthermore, the work by Niess et al. [53] indicate a link between rumination and self-compassion, which calls for more future research on how PI tools can be designed for self-compassion, i.e. to reduce negative processes like rumination.

Self-compassion is a concept supported by measurement instruments. Neff [47] created the Self-Compassion Scale (SCS) which measures the main components of self-compassion as mentioned previously. Raes et al. [58] constructed a short version of this scale, named the SCS-SF. Although some research has been conducted with the SCS, more is still needed (e.g. [50]). Neff et al. [50] used the SCS to measure participants self-compassion scores as well. The authors found that an increase in an individual's self-compassion was associated with an increase in other 'markers' of mental health, and that self-compassion functions as a protector against 'self-evaluative anxiety' in the context of considering one's own personal weaknesses [50]. Additionally, they found evidence to suggest that self-compassion has incremental validity as a construct, as its benefits 'could not be fully explained by lower levels of negative affectivity' [50, p. 150]. The measurement instruments for self-compassion present an opportunity for interaction design as they can be used to identify technologies which foster self-compassion.

2.3 Designing for Self-Compassion with Framing Effects

Although self-compassion is rarely—if ever—mentioned as a direct design goal for PI tools, there are some examples that show potential to design for compassion. For instance, Kocielnik et al. [38] designed an artefact to bridge the connection between reflection and well-being. Their artefact allowed users to reflect on the stressful factors in their lives [38], which poses the opportunity for users to become more compassionate with themselves. Furthermore, Epstein et al. [26] found that data visualisations from previously tracked data could give users more desire to get back to tracking after having lapsed, especially when a user had a stronger desire to return to tracking. This may show that data visualisations have the capacity to provide 'compassionate' support to a user.

Data elements such as textual information in the form of 'facts' (e.g. [23]) and labels (e.g. [41, 43]) help users evaluate their progress and interpret their data accurately in order to identify what actions to take, whilst low-level data could be more difficult for users to interpret. Low-level data representations, which tend to lack points of reference, are typically unhelpful for the interpretation of data as users are unable to identify contributing factors to their health behaviours [32]. Points of reference, like labels and factual information, can offer the user more opportunities for identifying contributing factors to their progress as these elements add more value to the data visualisation. This would imply that points of reference have the potential to trigger self-reflection processes in users. In other words, points of reference (e.g. [23]) and high-level data presentations (e.g. [32]) are key factors contributing to how users interpret PI data. Niess et al. [53] found that bar graphs also offered potential for self-reflection. Highlighted and labelled timelines (e.g. [38]) may be one of the most helpful data representations. Moreover, having a combination, and therefore a balance, in the type of data representations is important, as users tend to discover particular behavioural patterns more easily when given several 'perspectives' from the different kinds of data representations (e.g. [3]). Having a selection of different kinds of data visualisations is important for the discoverability of behavioural patterns. This finding tells us that several effective ways for displaying and presenting personal informatics exist, yet this makes it harder to concretely identify what particular data visualisation elements can present the user's 'successes' effectively and how to give users the ability to see the positive aspects in their data, especially under particular circumstances.

Visualising 'successes' and 'failures' in personal data in the form of framing techniques was not commonly studied in the field of PI. The light-based framing approach (colours green, yellow and red) (e.g. [41]) is an example of visualising the degree of success. Historically, framing was an important consideration in the context of persuasive technology [17, 19]. Past research suggests that positive framing techniques are preferred over negative ones (e.g. [18, 20, 26]). At the same time, positive and neutral framing techniques may not differ significantly in their emotional effects on the user (e.g. [18, 20]). Understanding the notion of framing and how it can be used to effectively enhance the experience of personal tracking is a key challenge as past work has shown that sensemaking in personal informatics often happens in the form of contextualising information. This was the case in a work-time tracking study by Epstein et al. [24]. Similarly, Bentley et al. [8] showed how multiple data sources about well-being offered benefits to users when framed in a context relevant way.

Across research in personal informatics, there is consensus that systems should provide context and reference points for the users' data in order to offer a better tracking experience. This was demonstrated in studying, inter alia, tracking habits [26], tracker metrics [11], and tracking fatigue [4]. Often, personal informatics is goal-driven and uses goals as points of reference [18]. While this offers a change-oriented experience, the focus on goals and, consequently, performance can also have negative implications [53]. Thus, the design challenge for future PI systems lies in determining how such contextual information could be delivered in a way that benefits the user. This, in turn, calls for the question of how to use framing to provide reference information to users of PI systems, as there is not enough evidence to suggest whether solely positive, neutral or even other types of framing approaches, have substantially different effects on (other) cognitive processes. Further, research regarding framing is typically focused on emotions and behaviour (e.g. [37]). Therefore, in this work, we aim to explore how data visualisation elements like baselines and framing techniques compare to one another in terms of their effect on self-reflection, rumination and self-compassion. As these aspects remain unexplored in PI research, our aim is to explore the role of framing and contextual information in PI experiences in a comparative study. For an initial exploration, we posit the simplest form of contextual information as a single reference data point and focus this work on such a form of context. We refer to this information as 'point of reference' for the remainder of this work.

2.4 Research Questions

It remains unclear what design choices should be made in PI tools in order to support self-compassion and prevent rumination. There is a research opportunity in understanding how to design PI systems that foster

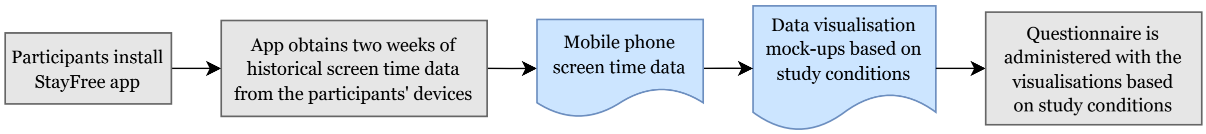

Fig. 1. Simplified flow chart of the steps that were taken, as well as the materials that were produced in the study. The grey-coloured blocks represent tasks that were either completed by the participants, the researchers, or the app. The blue-coloured blocks represent study materials, like data and mock-ups.

self-compassion, as it may be a valuable method for reducing potential rumination and improving the experience of PI tools [22]. Therefore, it is our goal to identify in what ways PI tools can improve the emotional experience, especially by focusing on the prevention of rumination, and triggering reflection and self-compassion. We explore how framing and a point of reference can influence these cognitive processes. We aim to answer the following research questions:

- **RQ1:** *In what ways do different kinds of framing techniques and points of reference in data visualisations affect self-reflection, rumination and self-compassion?*
- **RQ2:** *How can self-compassion contribute to the prevention of rumination in users who explore personal data?*

3 METHOD

We aimed to answer the research questions with an experimental study. This section describes the materials that were used, the recruitment of participants, and the study process. For the purpose of our work, we chose screen time data as it was easily obtainable in an online study and applicable to a broad group of participants. Further, screen time was shown to be a relevant source of data for past personal informatics studies [21, 60]. Additionally, screen time data is a relevant example to study as users often express dissatisfaction with their habits surrounding their smartphones, and understanding these behaviour patterns is relevant in managing them [45]. This requires a deeper understanding as to *when* and *why* users may start feeling dissatisfied with their habits [45]. Consequently, screen time tracking is an application area that can benefit from presenting data in a way that fosters positive cognitive processes, like reflection and self-compassion, and to inhibit negative ones, like rumination.

3.1 Study Design

To study user attitudes towards alternatives data visualisation in PI, we conducted a study where participants expressed their views on visualisations of their mobile device usage data. We obtained two weeks of historical data from the participants' devices. Then, we produced custom visualisations for each participant. Before the study, participants were provided an instructions document informing them of the study process, and a form asking for informed consent. The consent form included details on the anonymisation of their data. Figure 1 presents the structure of the study.

3.2 Conditions

In this study, we dedicated specific attention to exploring the interplay between various framing techniques and the inclusion of a point of reference in data visualisations, using screen time data as an example of a personal tracking experience. Our goal was to operationalise framing and contextual information so that we could conduct a comparative study. We operationalised the presence of contextual information as the addition of a single *point of reference*, the weekly average screen time value. We operationalised framing into negative, positive and neutral, as these framing are most often used in studies of affective framing across fields [35, 55]. By systematically varying these two components across visualisations, this study aims to dissect their effects on users' engagement with and understanding of their personal screen time data, providing critical insights for the design of more user-centred personal informatics tools.

We designed six different data visualisation mock-ups for each participant based on their personal screen time data in order to study the effects of different framing techniques in combination with a point of reference. All visualisations showed data from eleven days as this was the duration visualised by the *StayFree* app. We used bar graphs and reference lines as these visualisations are most commonly used in consumer-grade PI apps. Each mock-up given to the participant represented the same personal data with the same baseline. The baseline represented the participant's daily average screen time. Two mock-ups were framed positively by highlighting the days on which the participant had a daily screen time below their average in green, therefore highlighting the more 'successful' days (see Figures 2 and 3). Another two mock-ups were framed neutrally by colouring all days in grey (see Figures 4 and 5). Neither successes nor failures are highlighted with neutral framing. The last two were framed negatively by highlighting the days on which the participant had a daily screen time above their average in red, therefore highlighting the 'failures' (see Figures 6 and 7). Each framing technique had one visualisation with a point of reference and one without, resulting in two mock-up versions for each framing technique. A point of reference is added with the inclusion of labels to each bar in the graph, which corresponds to the total screen time for that day. All data visualisation mock-ups had the same graph type: the bar graph. This totals to six experimental conditions for the study:

- *PosReference:* Positive framing - point of reference (see Figure 2);
- *PosNo:* Positive framing - no point of reference (see Figure 3);
- *NeuReference:* Neutral framing - point of reference (see Figure 4);
- *NeuNo:* Neutral framing - no point of reference (see Figure 5);
- *NegReference:* Negative framing - point of reference (see Figure 6);
- *NegNo:* Negative framing - no point of reference (see Figure 7).

In this paper, we refer to the conditions with the abbreviations *PosReference*, *PosNo*, *NeuReference*, *NeuNo*, *NegReference* and *NegNo*.

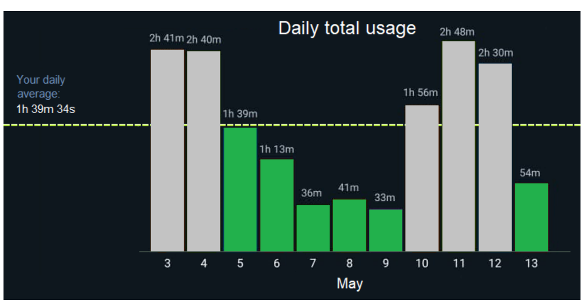

Fig. 2. *PosReference* - Positively framed bar chart with point of reference.

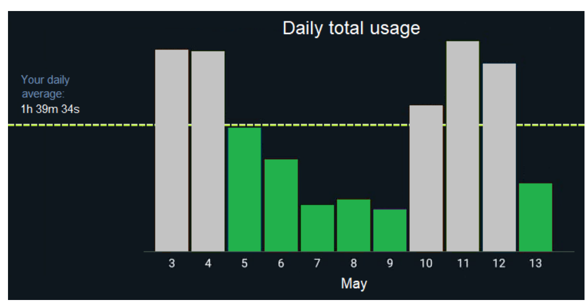

Fig. 3. *PosNo* - Positively framed bar chart without point of reference.

3.3 Measures

We gathered the data in this study through an online one-time questionnaire, designed with the *Qualtrics XM* platform. From this questionnaire, both quantitative and qualitative data is collected, as the questionnaire contains both Likert scale statements and open text fields for open questions.

In the first section of the questionnaire, participants were asked for their demographic data. This section of the questionnaire asks them about their age, gender, their experience with their personal screen time data, whether they have colour blindness or not and their highest completed education level.

The following sections of the questionnaire contained questions specific to the measures listed below.

3.3.1 Reflection and Rumination. In order to identify whether the personalised data visualisations triggered reflection or rumination, we used the Rumination–Reflection Questionnaire (RRQ) by Trapnell and Campbell [67] to construct questionnaire statements for the context of this study. To compare the reflection scores on the statements constructed from the RRQ, we also added questionnaire statements adapted from the Technology-Supported Reflection Inventory (TSRI) by Bentvelzen et al. [9], which is a scale developed with the purpose of comparing artefacts designed for reflection. Hence, we administered the following measures:

- Rumination Scale - 5 items on a 5-point Likert scale;
- Reflection Scale - 5 items on a 5-point Likert scale;
- Trait Reflection Scale - 12 items on a 5-point Likert scale;
- Trait Rumination Scale - 12 items on a 5-point Likert scale;
- TSRI scale - 9 items on a 7-point Likert scale.

The second section of the questionnaire contains statements regarding the participant's trait reflection and trait rumination. The participant's trait reflection and trait rumination are measured using the original statements from the RRQ by Trapnell and Campbell [67] with the so-called Trait Reflection Scale and the Trait Rumination Scale respectively. The Trait Rumination Scale and the Trait Reflection Scale both consist of 12 items on 5-point Likert scales. In total, this section contains 24 questions.

The next part of the questionnaire explored whether the data visualisation mock-ups triggered reflection or rumination. These reflection and rumination levels were measured with the help of two methods: the RRQ and the TSRI. Regarding the RRQ method, we used a modified version of the Reflection and Rumination Scales composed by Niess et al. [53], which consist of statements inspired by the RRQ from Trapnell and Campbell [67]. Both scales consist of 5 items on 5-point Likert scales. For the TSRI method, we used modified statements from the TSRI in order to evaluate whether the data visualisation supports reflection or not. The TSRI scale consists of 9 items on 7-point Likert scales.

The participant rates the statements on the scales from both methods for every single condition. Therefore, this section of the questionnaire contained $(9 * 6) + (5 * 6) + (5 * 6) = 54 + 30 + 30 = 114$ questions.

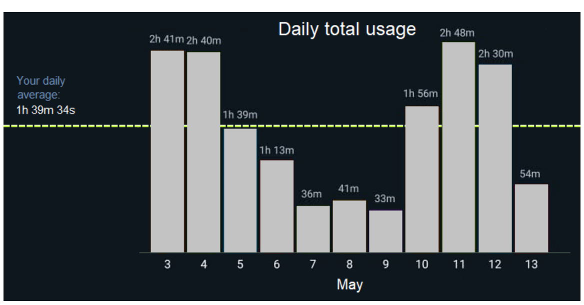

Fig. 4. *NeuReference* - Neutrally framed bar chart with a point of reference.

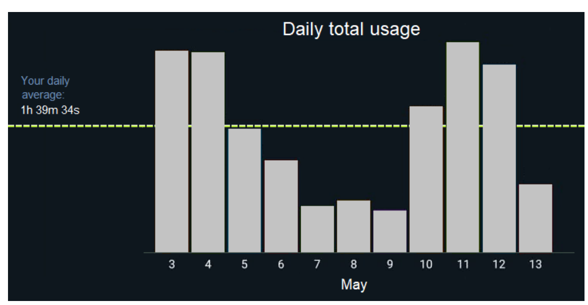

Fig. 5. *NeuNo* - Neutrally framed bar chart without a point of reference.

3.3.2 Self-Compassion. In order to identify whether the personalised data visualisations triggered self-compassion, we used the SCS-SF by Raes et al. [58] to construct questionnaire statements for the context of this study. We administered the following measures:

- Self-Compassion Scale - 5 items on a 5-point Likert scale;
- Trait Self-Compassion Scale - 12 items on a 5-point Likert scale.

Next, the questionnaire contained statements regarding the participant's trait self-compassion. The participant's trait self-compassion was measured using the original statements from the SCS-SF by Raes et al. [58] with the Trait Self-Compassion Scale. The Trait Self-Compassion Scale consists of 12 items on a 5-point Likert scale. The next part of the questionnaire administered the degree to which self-compassion is triggered by the data visualisation in the participant. The self-compassion level was measured with the help of the Self-Compassion Scale, which is a scale that consists of 5 statements inspired by the original statements of the SCS-SF by Raes et al. [58]. Hence, the Self-Compassion Scale is a scale consisting of 5 items that are rated on 5-point Likert scales.

The participant first rated the statements from the Trait Self-Compassion Scale so that their trait self-compassion can be used as a covariate in the statistical analyses. Then, the participant rated the statements on the Self-Compassion Scale for all six conditions. Therefore, this section of the questionnaire contains 12 + (6 * 5) = 12 + 30 = 42 questions.

The exact statements used as measures in the study questionnaire are provided in the auxiliary material.

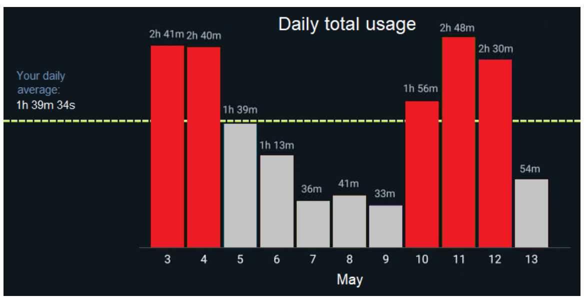

Fig. 6. *NegReference* - Negatively framed bar chart with point of reference.

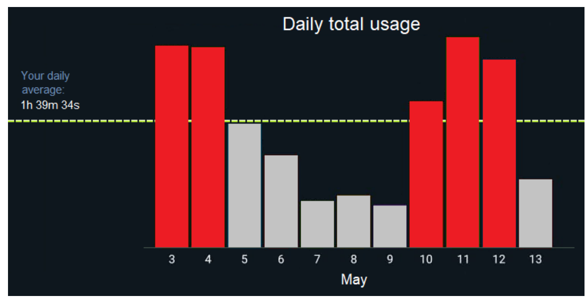

Fig. 7. *NegNo* - Negatively framed bar chart without point of reference.

3.4 Open Questions

The last section of the questionnaire contained three questions with open text fields in order to collect qualitative data. The first question asked the participant how they would identify their goals for their screen time data. The second question asked the participant to describe their idea of a 'success' and a 'failure' when looking at their screen time data. In the last question, we ask participants how looking at particular data visualisations made them feel. By analysing the participants' answers to these questions, we investigated in what ways the representations of successes and failures in the mock-ups may have influenced the results.

3.5 Questionnaire Administration

Here, we describe how we obtained participant data and later administered the online questionnaire, using visualisations based on the participants' device usage statistics.

3.5.1 Obtaining User Data. The participants were asked to download the *StayFree* (see: https://stayfreeapps.com/) mobile app so that they could send us their screen time data. Participants were given the full instructions for installing the mobile app and sending the data in the accompanying consent form. After completing all the steps successfully, the participants shared an image of their screen time either over text messaging or email. For this study, the use of screen time data was used instead of, e.g., fitness data collected from fitness trackers, as most people nowadays have a smartphone with automatic screen time tracking function, whilst a limited group of people collect their fitness data. Furthermore, screen time tracking is relevant across a variety of different user groups due to its impact on health, productivity, and mindfulness, amongst other aspects.

Afterwards, within five days from receiving the user data, we designed six personalised data visualisations for each participant based on their screen time data. Using personal data made the data visualisations more relevant to the participant, potentially increasing the participant's interest when taking the questionnaire. This was taken into consideration due to the length of the questionnaire (approximately 30 minutes). Engaging with their own data was more likely to ensure quality answers than working with hypothetical data, which is a recognised challenges in PI studies [53]. All participants provided their data within two weeks from recruitment.

3.5.2 Administering the Questionnaire. After all the participants had sent their screen time data, the participants were contacted via text messaging or email and sent the study questionnaire with their personalised data visualisations. Before sending out the questionnaire to all participants, we did a trial session with one participant to pilot the questionnaire and ensure the questionnaire did not contain any inconsistencies or vague question formulations. All participants were asked to complete the questionnaire within a period of three weeks.

The participants were asked hypothetical questions, in which they had to answer whether they would reflect or ruminate with the given data visualisation. They were also asked regular questions with regards to how they felt about the shown data visualisation with scales that focus on feelings of self-compassion. We opted for this type of study, as it offers the advantages of both experimental studies and applied studies: high internal validity and high external validity respectively [2].

The participants were asked to answer the same questions for all conditions. Therefore, this study had a repeated-measures (within-subjects) design.

3.6 Participants

To recruit participants, we used convenience, as this study required the users to have an Android mobile phone (at the time of the study, the *StayFree* mobile app was solely available for Android). Snowball sampling was used as well.

The sample consisted solely of adults based in the European Union. We recruited 18 participants (n = 18), with 13 females (72.22%) and 5 males (27.78%). The participants were aged from 19 to 26 years old (M = 22.38, SD = 1.79). When asked how often they engaged in exploring personal data, including screen time data, 9 participants (50%) never spent time doing so, 5 (27.78%) once a month, 2 (11.11%) more than once a month and another 2 (11.11%) explored their data every week. All participants reported that they did not have any form of colour blindness. Table 1 provides an overview of the details about the participants.

3.7 Data Analysis

3.7.1 Quantitative Analysis. To analyse the quantitative data, we used ANOVA-type procedures appropriate to the specific dependent variables, with ANCOVAs to control for covariance based on past results when available. Post-hoc analysis was conducted using Tukey's HSD. The repeated-measures model specifications used assured that we compensated for the within-subjects nature of our study.

3.7.2 Qualitative Analysis. To analyse the qualitative data, we used open coding and the pragmatic approach to qualitative analysis as described by Blandford et al. [13]. The first author open coded all the answers from the participants. Afterwards, in an iterative discussion, authors grouped and revised the codes using affinity diagramming to build an understanding of the data corpus. Eventually, we constructed five themes which comprehensively describe the contents of the qualitative data sources from the questionnaires.

Table 1. Details of the participants in the study, including their participant ID (PID), gender, age and highest completed education level.

PID	Gender	Age	Education
P1	Male	22	High school
P2	Female	21	High school
P3	Male	22	High school
P4	Female	21	Bachelor
P5	Female	21	High school
P6	Female	22	High school
P7	Female	21	High school
P8	Female	21	High school
P9	Female	25	Bachelor
P10	Male	22	Bachelor
P11	Female	22	Bachelor
P12	Female	22	High school
P13	Male	26	Bachelor
P14	Female	19	High school
P15	Male	25	Master
P16	Female	23	High school
P17	Female	24	High school
P18	Female	24	Bachelor

4 RESULTS

Here, we present the results of the study. First, this section provides an overview of the quantitative results. This is followed by the findings based on the qualitative analysis of the open questions. In the auxiliary material, we provide the raw quantitative data from all participants.

4.1 Reflection Scale

To investigate the effect of the type of data visualisation on the Reflection Scale score, we conducted a two-way ANCOVA. We controlled for the participants' reflection trait as measured on the Trait Reflection Scale. The framing technique had a significant effect on the reflection levels, $F(2, 102) = 10.328$, $p < 0.001$. There was a significant effect of the covariate, the Trait Reflection, F(1, 102) = 15.787, p < 0.001. The in- or exclusion of a point of reference had no significant effect on the reflection levels, F(1, 102) = 2.395, p = 0.125. There was no statistically significant interaction between the framing technique and the in- or exclusion of a point of reference on the reflection score, whilst controlling for trait reflection, F(5, 102) = 0.246, p = 0.782.

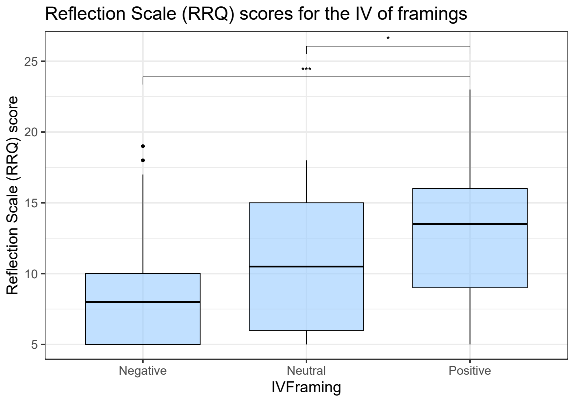

Fig. 8. Boxplots of the scores on the Reflection Scale for the IV concerning the framings, including significant pairwise differences.

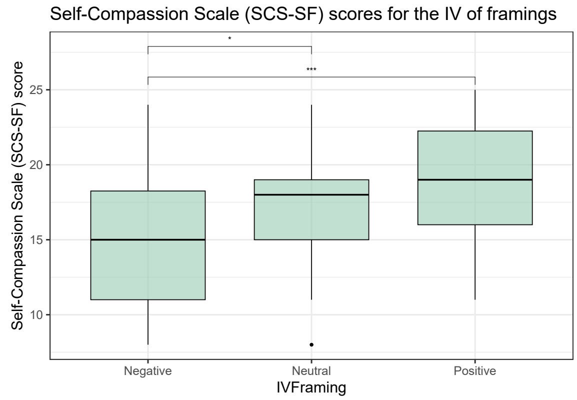

Fig. 9. Boxplots of the scores on the Self-Compassion Scale for the IV concerning the framings, including significant pairwise differences.

In order to investigate significant pairwise differences between the six conditions on the Reflection Scale, we conducted a post-hoc test with Tukey's HSD. This revealed significant differences between the condition pairs *PosReference - NegNo* with $p < 0.01$, and *PosNo - NegNo*, *PosReference - NeuNo* and *PosReference - NegReference* with $p < 0.05$. Additionally, this revealed significant differences between framing techniques **positive** and **negative** with $p < 0.001$, and **positive** and **neutral** with $p < 0.05$. Figure 8 shows the scores on the Reflection Scale for the framing techniques.

4.2 Self-Compassion Scale

For the investigation of the effect of the type of data visualisations on the Self-Compassion Scale score, we conducted a two-way ANCOVA, controlling for the participants' self-compassion trait as measured on the Trait Self-Compassion Scale. The framing technique had a significant effect on the self-compassion levels, $F(2, 102) = 11.948$, $p < 0.001$. Also, the covariate Trait Self-Compassion had a significant effect, $p < 0.001$. The in- or exclusion of a point of reference had no significant effect on the self-compassion levels, $F(1, 102) = 0.099$, $p = 0.753$. No statistically significant interaction was found between the framing technique and the in- or exclusion of a point of reference on the self-compassion score, whilst controlling for trait self-compassion, $F(5, 102) = 0.109$, $p = 0.897$.

Post-hoc testing with Tukey's HSD revealed significant differences between the condition pairs *PosReference - NegReference* and *NegNo - PosReference* with $p < 0.05$, and *PosNo - NegReference* and *PosNo - NegNo* with $p < 0.01$. Additionally, post-hoc testing revealed significant differences between framing techniques **neutral** and **negative** with $p < 0.05$, and **positive** and **negative** with $p < 0.001$. Figure 9 shows the scores on the Self-Compassion Scale for the three framing techniques.

4.3 TSRI

We conducted a two-way ANOVA to investigate the effect of the type of data visualisation on the TSRI scale score. The framing technique had a significant effect on the reflection levels, $F(2, 102) = 4.769$, $p < 0.05$. Furthermore, the in- or exclusion of a point of reference had a significant effect on the reflection levels, $F(1, 102) = 5.716$, $p < 0.05$. There was no statistically significant interaction between the framing technique and the in- or exclusion of a point of reference on the reflection score, $F(5, 102) = 0.181$, $p = 0.835$.

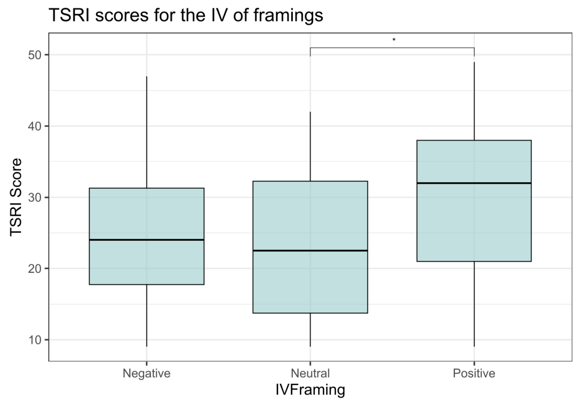

Fig. 10. Boxplots of the scores on the TSRI Scale for the IV concerning the framings, including significant pairwise differences.

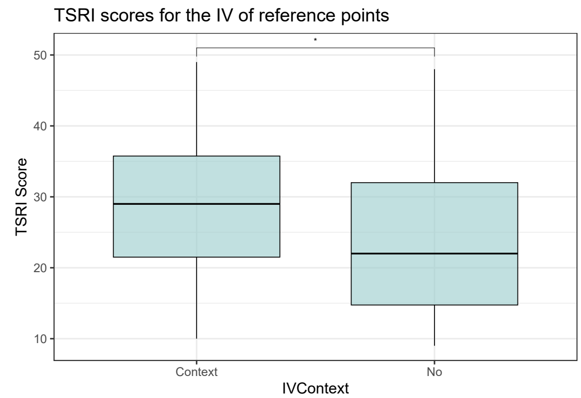

Fig. 11. Boxplots of the scores on the TSRI Scale for the IV concerning the points of reference, including significant pairwise differences.

A post-hoc test with Tukey's HSD for the TSRI scale revealed significant pairwise differences between the condition pair *NeuNo - PosReference* with $p < 0.01$. Subsequently, post-hoc testing revealed significant differences between framing techniques **positive** and **neutral** with $p < 0.05$ and the **in- or exclusion of a point of reference** with $p < 0.05$. Figure 10 shows the scores on the TSRI Scale for the three framing techniques. Figure 11 displays the significant difference for the point of reference.

4.4 Rumination Scale

To investigate the effect of the type of data visualisation on the Reflection Scale score, we conducted a two-way ANCOVA. We controlled for the participants' rumination trait as measured on the Trait Rumination Scale. The framing technique had a significant effect on the rumination levels, $F(2, 102) = 16.814$, $p < 0.001$. Additionally, the in- or exclusion of a point of reference had a significant effect, $F(1, 102) = 4.793$, $p < 0.05$. There was no significant effect of the covariate, the Trait Rumination, $F(1, 102) = 1.310$, $p = 0.255$. There was no statistically significant interaction between the framing technique and the in- or exclusion of a point of reference on the rumination score, whilst controlling for trait rumination, $F(5, 102) = 0.389$, $p = 0.679$.

For the Rumination Scale, we also tested for pairwise differences between the six conditions. The post-hoc with Tukey's HSD revealed significant differences between the condition pairs *NeuReference - NegReference* and *PosNo - NegNo* with $p < 0.01$, *NeuReference - NegNo* with $p < 0.05$, *NeuNo - NegNo*, *PosNo - NegReference* with $p < 0.001$, and *NeuNo - NegReference* with $p < 0.001$. Subsequently, post-hoc testing revealed significant differences between framing techniques **neutral** and **negative** with $p < 0.001$, and **positive** and **negative** with $p < 0.001$. Additionally, post-hoc testing revealed significant differences between the **in- or exclusion of a point of reference** with $p < 0.05$. Figure 12 shows the scores on the Rumination Scale for the different framing techniques. Figure 13 displays the significant difference between the exclusion and inclusion of a point of reference.

4.5 Qualitative Results

While some participants mentioned that they did not have any particular goal or feeling towards the data visualisations, many participants reflected on their experiences in more detail. Based on our analysis, five themes have been identified: *Comparison, Insight, Thresholds, Context,* and *Framing*. In the following, we describe the identified themes in more detail, illustrating them with excerpts from the data collection.

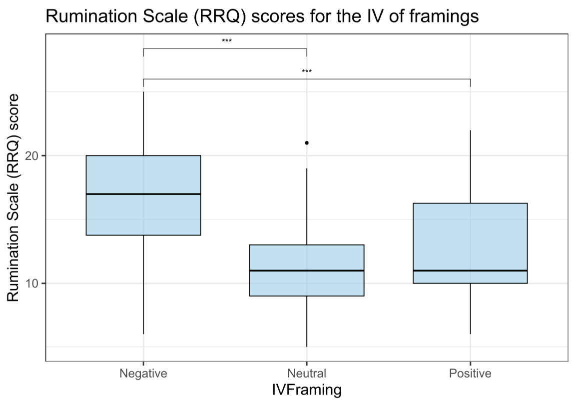

Fig. 12. Boxplots of the scores on the Rumination Scale for the IV concerning the framings, including significant pairwise differences.

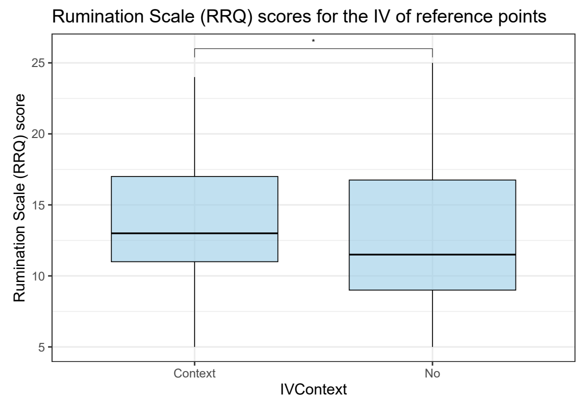

Fig. 13. Boxplots of the scores on the Rumination Scale for the IV concerning the points of reference, including significant pairwise differences.

4.5.1 Comparison. The first theme derived from our data focuses on the potential and the relevance of drawing comparisons between your own data and other data. The comparisons can either be drawn between individual data or can go beyond that. For instance, one participant remarked the potential of drawing comparisons between their own data and recognised benchmarks in the area of health: *'[My goal is] to try and compare my own data with the approved / healthy standards.' (P3)*. We observed that participants often based their interpretations of data around comparison and aimed to identify features of the data representation which enabled that. Users used different time horizons to compare data points, analysing the data across days, weeks or the full visualised day range.

4.5.2 Insight. The second theme describes one essential way of how participants describe that their data should be used, namely to help them learn more about themselves in a non-judgemental way. Participants do not necessarily want to set a specific goal with regards to their (screen time) data, rather they were more interested in gaining insight from the data visualisations.

This is nicely illustrated by the comment of one participant that emphasised the need to make their own assessment of how they are doing, instead of allowing the technology to make this judgement for them: *'(...) to provide myself with insight, not letting it [i.e. the technology] tell me how "good" or "bad" I'm doing something. I decide that myself.' (P11)*

4.5.3 Thresholds. Along similar lines, this theme focuses on self-determined decisions. Participant voiced that they want to have the power to set their own goals, determine thresholds, decide for a specific framing and take responsibility for selecting specific parts of their screen time experience they want to reduce. Furthermore, they note that the average used in the study may also be considered a sufficient starting point to define successes and failures.

On the other hand, participants also discussed that an average is not an ideal measure for success of failure, because an average is a dynamic metric. It can change every single week or even every day depending on your performance from the previous week or day. Several participants preferred setting specific limits or thresholds over averages:

- *'I think you can only talk about success or failure when you have set a personal goal. [...] If your goal is to get below average, the goal constantly changes because your average will go down as well. [...] I think you should set your goal at your personal threshold.' (P17)*
- *'I don't think average is a good way to measure success, as you could also have the same screen time everyday, but that doesn't mean it's either good or bad. [...] Starting with average might be good as an indication, but after that you could maybe lower the bar slowly.' (P9)*

4.5.4 Context. This theme focuses on the idea that quality of screen time can be more important than the actual amount of screen time. For example, some mobile applications could be considered useful, or even necessary for essential tasks. One participant mentioned that using her phone is actually necessary for school at times: *'[...] a lot of my screen time is actually also because of school, so it would be stupid to see that as a negative thing.' (P2)*

Another participant denoted that knowing when to stop using their phone in itself could be a success: *'A success in this case would be to stop looking at the phone when looking at the phone becomes detrimental to other aspects in your life.' (P13).* This pinpoints the idea that successes or failures with regards to the representation of data does not necessarily have to be defined in terms of numbers. Some users may find context and quality more important.

4.5.5 Framing. The emotional effect that different kinds of data representations had on participants differed which falls under the umbrella of the theme *framing*. However, we did identify one common aspect: emphasising failures, especially with a distinct colour like red, induced negative emotions. Many participants declared feeling negative emotions when being confronted with those mock-ups. Common emotional responses to the conditions *NegContext* and *NegNo* were sadness, annoyance, discouragement and frustration. One participant elaborated on the fact that these mock-ups would make it difficult to make changes to their behaviour as they would want to avoid looking at the graph: *'The failures graph made me not want to look at it. At first glance you only see the bad days, that way I can only think of that [...]. It would make it difficult to think of ways to change.' (P12)*

Furthermore, very few participants mentioned the positive framing technique. This could imply that the emotions induced by the negative framing technique were stronger compared to those induced by the positive framing technique. This could explain why more participants expressed their negative emotions, rather than their positive ones.

5 DISCUSSION

This study's focus on screen time data visualisation raises a compelling question about the adaptability of the findings in different PI contexts, such as health or fitness tracking. The choice of screen time as the context for this study stems from its pervasive relevance in modern digital life, but the implications of our findings extend beyond this domain. Designing data visualisations with self-compassion in mind can be a promising approach across various PI domains, where negative self-comparisons and rumination [53] might be triggered by data representations. For instance, in health and fitness tracking, positive or neutral framing techniques might help users to engage with their data in a way that encourages self-understanding rather than self-judgement [59], thereby promoting emotional well-being. It suggests a paradigm shift in PI tool design, moving from an exclusive focus on performance metrics towards incorporating users' emotional and mental well-being as primary considerations.

5.1 Understanding Framing and Points of Reference

In this research, we explored how the framing (positive, neutral, negative) of visualisations for screen time data affects reflection, rumination and self-compassion. Furthermore, we inquired how a point of reference provided in these visualisations can potentially support positive PI experiences. Having explored the effect of the data visualisations, we observed that the framing technique had a significant effect on a person's reflection, rumination and self-compassion levels. Framing technique had a significant effect on the DVs as measured on the Reflection, TSRI, Rumination and Self-Compassion scales. With the pairwise comparisons from post-hoc testing and the qualitative analysis, we were able to find in what ways the framing techniques had a significant effect on the DVs. We found that positive framing had a significant effect on the reflection score for the Reflection Scale over neutral and negative framing. For the reflection score on the TSRI scale, positive framing had a significant effect over neutral framing. As shown in the boxplots from Figures 8 and 10, positive framing culminates higher reflection scores (**RQ1**). Subsequently, negative framing had a significant effect on the rumination score for the Rumination Scale over positive and neutral framing. Consequently, Figure 12 shows that negative framing leads to higher rumination scores (**RQ1**). This finding is in line with theories of reflection and rumination, since reflection is considered a positive process [67] and rumination a negative one [46, 64]. Furthermore, negative framing had a significant effect on the self-compassion score on the Self-Compassion Scale over positive and neutral framing. Figure 9 shows that negative framing results in noticeably lower self-compassion scores (**RQ1**, **RQ2**). This finding is comparable to previous findings as self-compassion is negatively associated with negative processes and feelings [47, 51]. This negative association seems to be illustrated with the effect that negative framing has on self-compassion. Similarly, the qualitative analysis revealed that the negative framing technique seems to result in a strong negative emotional response in users. In conclusion, positive framing increases self-reflection, negative framing increases rumination but decreases self-compassion. This finding shows the potential to design with positive or non-negative framing techniques and self-compassion in mind in order to trigger self-reflection and discourage rumination (**RQ1**, **RQ2**). To design for self-compassion, PI tools should avoid negative framings and instead opt for positive or neutral framing techniques (**RQ2**).

The in- or exclusion of a point of reference had a significant effect solely on the TSRI and Rumination scales. Interestingly, a point of reference seems to result in higher reflection scores compared to the exclusion of context (**RQ1**). In section 2.3, we hypothesised that context has the potential to trigger self-reflection in users, yet comparable to the quantitative results, the qualitative results revealed that attitudes towards the a point of reference were mixed. This could imply that the effect of context is more personal, rather than universal. Therefore, future PI tools should have customisation options so users can decide whether they want to have context visualised in their data in order to prevent negative emotions. Another suggestion is to develop future PI tools that can make automatic adaptations to the data visualisation depending on, e.g., the user's behaviour.

We note that, fundamentally, framing is a type of cognitive bias [68]. Consequently, using framing in PI systems implies a certain less of objectivity, which, in turn, may become a key design consideration. Design solutions where objectivity is a design goal may attempt to avoid the bias, e.g., by presenting data in both positive and negative forms. Previous studies have shown that both positive [18] and negative framing [37] can be effective in supporting users with behaviour change. In contrast, here, we investigated how PI systems may support reflection, which is seen as a key way for PI to empower users to make informed decisions about their well-being [12]. Our results showcase that there is room for using framing in the context of reflection and positive framing is preferred. Future studies should investigate if that is also the case across different PI systems and contexts.

5.2 The Role of Self-Compassion

In light of our findings, the integration of self-compassion as an aspect of Personal Informatics (PI) tool design emerges as a viable alternative to performance-oriented tracking, which is most often seen in commercial products [12, 65]. Through our research, we demonstrated that self-compassion can affect how individuals perceive and react to data about themselves. The demonstrated impact of framing, particularly positive and neutral techniques, upon self-reflection and self-compassion scores, suggests that designing with empathy and understanding can foster positive interactions with PI tools. Interestingly, this perspective diverges from conventional design paradigms that predominantly focus on performance metrics, thus offering an approach that embraces emotional well-being as a key design goal for PI. Self-compassion can potentially contribute to the reduction of self-critical rumination, an unintended consequence of performance-centric PI tools. This integration of self-compassion in designing PI provides a valuable avenue for future PI tools that prioritise user emotional health and self-understanding. The possibility of customising data visualisation based on the user's behaviour or preferences further underscores this approach, marking a shift from one-size-fits-all solutions to more personalised tools. This study highlights the transformative potential of self-compassion in PI tool design, opening new directions for future research and development in this field.

Acknowledging the lab-study nature of our research, wherein participants viewed different visualisations consecutively, we recognise that this design may have primed participants to notice and comment on the differences between conditions. As an initial exploration into self-compassion in PI, this study aimed to capture participants' initial reactions to different data visualisations. However, future work could extend this research to a between-subjects evaluation and/or a field study to explore how these visualisations affect users in more ecologically rigorous, long-term settings. Moreover, our study opens the door for exploring additional forms of positive reinforcement in PI contexts. For instance, integrating personalised affirmations, goal-setting that emphasises self-care [59], or collaborative features that foster social support [57] could be promising avenues for fostering positive, compassionate engagement with PI tools. These contributions offer the community a roadmap for integrating self-compassion into the design of PI tools in a manner that aligns with the needs and emotional well-being of end-users.

5.3 Ways Forward

In this paper, we have identified several directions for future research. The first research opportunity is to control for character traits. In this paper, we controlled for a person's natural tendency to ruminate or self-reflect, and for the person's self-compassion trait. In PI research, a user's character traits are typically not taken into consideration when measuring or evaluating the DV. This finding shows that this is a potential research gap in the field of HCI and personal informatics. Future research should aim to fill this research gap and determine to what extent and in what ways character traits could affect a person's perception of their data.

Second, in line with past research [7, 12], our work shows that more research is needed in order to identify and develop research methods for evaluating and measuring reflection, especially with regards to the evaluation of artefacts. The literature review revealed that papers may have different takes on what reflection means or entails, therefore choosing different research methods for evaluating this process in users when they explore research artefacts. This may complicate comparing the results from various studies and identifying concrete design choices for PI tools that explain the artefacts' effects on the user. Future research should aim to explore effective ways in which to study reflection and rumination in the context of PI, allowing for effective system comparison.

Third, future research should consider taking participants' personal goals and definitions of success and failure into account. Our study confirms the need for considering the complexity of tracking goals [54] in designing data representations for PI. For example, several participants stated that successes and failures could depend on the context of the data, whilst others stated that setting a rigid personal goal achievement threshold could correspond to their goals. These differences in goal setting and perceptions of success could influence one's experience with a data visualisation, hence their engagement with the data as well. This consideration for personal differences also applies to the addition of a point of reference in data visualisations, since we found that context resulted in mixed reactions from participants. Therefore, future research should take these personal differences into account when exploring the effects of a data visualisation.

Lastly, we aimed to explore whether a person's self-compassion was affected by the different kinds of data representations or not. Although the study showed that negative framings should be avoided when designing for self-compassion, more research is needed in order to identify what design elements are most effective with the aim to properly foster self-compassion, including the self-compassion plays in discouraging rumination, especially in the case of goal failure.

5.4 Limitations

We acknowledge several limitations of our approach. Parts of this research were conducted during the Covid-19 pandemic. Consequently, the study could not benefit from in-person participant interactions. Instead, contact with the participants was maintained via email and personal messaging (e.g. Microsoft Teams). Recruitment of participants was mostly done online, which ultimately may have resulted in less participants taking part in the study than preferred. One limitation to this research, due to this non-personal type of communication exclusively, is the possibility that participants might have felt less engaged with the study, which could have altered the results. Also, there is the possibility that participants might have felt less engaged with the mock-ups that were given in the questionnaire, because not all participants had been actively tracking their screen time data. We tried to minimise this lack of engagement with the mock-ups by having the participants send over their personal screen time data first. This way, we could design mock-ups that presented their own data.

Further, we reflect on the the method of recruiting the participants through convenience sampling and snowball sampling. The recruitment of participants was not random, because we contacted friends and acquaintances to be potential participants in study. Subsequently, we asked several participants to recruit others from their inner circles to take part in the study. Therefore, we used a non-probability sampling method to recruit participants for the study. Due to the recruitment methods used for obtaining this sample, the results could be biased and have a lower external validity than preferred, because the sample might not be fully representative for the population of smartphone users that regularly evaluate their screen time. In the recruitment process, we tried to minimise this bias by recruiting participants of different ages, genders and education levels. However, we recognise that the sample consist primarily of young adults.

We acknowledge that our within-subjects design, in which participants were exposed to multiple conditions, might have primed participants to notice the differences between these conditions, potentially influencing their responses and the comparative evaluations they made. This design choice, while allowing for controlled and direct comparisons across conditions, could introduce carryover effects, which may have impacted the later measures in the survey, especially given the non-randomised order of the questions despite the conditions being order-balanced. Our particular study design may have also induced participant fatigue. Due to the combination of using several different scales and having a total of 6 conditions, the survey required a considerable amount of time to complete. This could have especially influenced the measures which appeared later in the survey Response fatigue could have altered the focus and motivation of the participants to fill in the questionnaire.

6 CONCLUSION

This paper aimed to evaluate in what ways different kinds of framing techniques and the addition of a point of reference in the representation of personal data can contribute to the design, specifically the interface design, of PI tools. We conducted a within-subject online study in which we evaluated how the dependent variables reflection, rumination and self-compassion levels of the participant were influenced by the two independent variables: the framing technique and the point of reference. From the statistical analyses, we found that the framing technique had a significant effect on the cognitive processes of the user. Positive framing showed a significant increase in self-reflection, whilst negative framing showed a significant increase in rumination and decrease in self-compassion. On the other hand, the effects of context differed per scale. More than that, the quantitative analysis revealed mixed results regarding the inclusion of baselines.

The qualitative analysis revealed that the negative framing technique induced a strong negative emotional response in the participants. Participants experienced a wide range of negative emotions, including sadness, annoyance, discouragement and frustration. This paper contributes to the field of personal informatics by providing insights into the design of future PI tools with the aim to foster self-reflection and self-compassion, and to discourage rumination in users. We hope to inspire future research on the design of data visualisations with the intent to minimise the effects of negative cognitive and emotional processes.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council, award number 2022-03196.

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), 165–178. https://doi.org/10.1037/0022-3514.92.2.165

[2] 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. https://doi.org/10.1177/1094428114547952

[3] Fereshteh Amini, Khalad Hasan, Andrea Bunt, and Pourang Irani. 2017. Data representations for in-situ exploration of health and fitness data. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare* (Barcelona, Spain) *(PervasiveHealth 2017)*. 163–172. https://doi.org/10.1145/3154862.3154879

[4] Amid Ayobi, Paul Marshall, and Anna L Cox. 2016. Reflections on 5 years of personal informatics: Rising concerns and emerging directions. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, CA, USA) *(CHI EA 2016)*. 2774–2781. https://doi.org/10.1145/2851581.2892406

[5] Ruth A Baer. 2003. Mindfulness training as a clinical intervention: a conceptual and empirical review. *Clinical Psychology: Science and Practice* 10, 2 (2003), 125–143. https://doi.org/10.1093/clipsy.bpg015

[6] Laura K Barnard and John F Curry. 2011. Self-compassion: Conceptualizations, correlates, & interventions. *Review of General Psychology* 15, 4 (2011), 289–303. https://doi.org/10.1037/a0025754

[7] Eric PS 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* (Crossings, Seoul, Korea) *(CHI 2015)*. 585–594. https://doi.org/10.1145/2702123.2702234

[8] 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 Transactions on Computer-Human Interaction (TOCHI)* 20, 5, Article 30 (11 2013), 27 pages. https://doi.org/10.1145/2503823

[9] Marit Bentvelzen, Jasmin Niess, Mikołaj P Woźniak, and Paweł W Woźniak. 2021. The development and validation of the Technology-Supported Reflection Inventory. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI 2021)*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3411764.3445673

[10] Marit Bentvelzen, Jasmin Niess, and Paweł W Woźniak. 2021. The Technology-Mediated Reflection Model: Barriers and assistance in data-driven reflection. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI 2021)*. ACM, New York, NY, USA, Article 246, 12 pages. https://doi.org/10.1145/3411764.3445505

[11] Marit Bentvelzen, Jasmin Niess, and Paweł W Woźniak. 2023. Designing reflective derived metrics for fitness trackers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4, Article 158 (12 2023), 19 pages. https://doi.org/10.1145/3569475

[12] Marit Bentvelzen, Paweł W Woźniak, Pia SF Herbes, Evropi Stefanidi, and Jasmin Niess. 2022. Revisiting reflection in HCI: Four design resources for technologies that support reflection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1, Article 2 (3 2022), 27 pages. https://doi.org/10.1145/3517233

[13] Ann Blandford, Dominic Furniss, and Stephann Makri. 2016. Qualitative HCI research: Going behind the scenes. *Synthesis Lectures on Human-Centered Informatics* 9, 1 (4 2016), 115 pages. https://doi.org/10.2200/S00706ED1V01Y201602HCI034

[14] Sidney J Blatt. 1995. Representational structures in psychopathology. *Cicchetti, D and Toth, S (Eds.), Rochester Symposium on Developmental Psychopathology: Emotion, Cognition, and Representation* 6, 1–34.

[15] Fadi Botros, Charles Perin, Bon Adriel Aseniero, and Sheelagh Carpendale. 2016. Go and Grow: Mapping personal data to a living plant. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Bari, Italy) *(AVI 2016)*. 112–119. https://doi.org/10.1145/2909132.2909267

[16] Juliana G Breines and Serena Chen. 2012. Self-compassion increases self-improvement motivation. *Personality and Social Psychology Bulletin* 38, 9 (2012), 1133–1143. https://doi.org/10.1177/0146167212445599

[17] Eun Kyoung Choe, Jaeyeon Jung, Bongshin Lee, and Kristie Fisher. 2013. Nudging people away from privacy-invasive mobile apps through visual framing. In *Human-Computer Interaction–INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part III 14*. Springer, 74–91. https://doi.org/10.1007/978-3-642-40477-1

[18] Eun Kyoung Choe, Bongshin Lee, Sean Munson, Wanda Pratt, and Julie A Kientz. 2013. Persuasive performance feedback: The effect of framing on self-efficacy. In *AMIA Annual Symposium Proceedings*, Vol. 2013. American Medical Informatics Association, 825–833.

[19] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding self-reflection: How people reflect on personal data through visual data exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare* (Barcelona, Spain) *(PervasiveHealth 2017)*. 173–182. https://doi.org/10.1145/3154862.3154881

[20] 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* (Seoul, Korea) *(UbiComp 2008)*. 54–63. https://doi.org/10.1145/1409635.1409644

[21] Nediyana Daskalova, Karthik Desingh, Alexandra Papoutsaki, Diane Schulze, Han Sha, and Jeff Huang. 2017. Lessons learned from two cohorts of personal informatics self-experiments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3, Article 46 (9 2017), 22 pages. https://doi.org/10.1145/3130911

[22] Elizabeth Victoria Eikey, Clara Marques Caldeira, Mayara Costa Figueiredo, Yunan Chen, Jessica L Borelli, Melissa Mazmanian, and Kai Zheng. 2021. Beyond self-reflection: Introducing the concept of rumination in personal informatics. *Personal and Ubiquitous Computing* 25, 3 (2021), 601–616. https://doi.org/10.1007/s00779-021-01573-w

[23] 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* (Vancouver, BC, Canada) *(DIS 2014)*. 667–676. https://doi.org/10.1145/2598510.2598558

[24] Daniel A Epstein, Daniel Avrahami, and Jacob T Biehl. 2016. Taking 5: Work-breaks, productivity, and opportunities for personal informatics for knowledge workers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, CA, USA) *(CHI 2016)*. 673–684. https://doi.org/10.1145/2858036.2858066

[25] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4, Article 126 (12 2020), 38 pages. https://doi.org/10.1145/3432231

[26] Daniel A Epstein, Jennifer H Kang, Laura R Pina, James Fogarty, and Sean A Munson. 2016. Reconsidering the device in the drawer: Lapses as a design opportunity in personal informatics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp 2016)*. 829–840. https://doi.org/10.1145/2971648.2971656

[27] 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* (Osaka, Japan) *(UbiComp 2015)*. 731–742. https://doi.org/10.1145/2750858.2804250

[28] Clayton Feustel, Shyamak Aggarwal, Bongshin Lee, and Lauren Wilcox. 2018. People like me: Designing for reflection on aggregate cohort data in personal informatics systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3, Article 107 (9 2018), 21 pages. https://doi.org/10.1145/3264917

[29] Paul Gilbert. 2005. Compassion and cruelty: A biopsychosocial approach. In *Compassion: Conceptualisations, Research and Use in Psychotherapy*. London: Routledge, 9–74. https://doi.org/10.4324/9780203003459

[30] 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 (WWW 2019)*. 571–582. https://doi.org/10.1145/3308558.3313432

[31] Rúben Gouveia, Evangelos Karapanos, and Marc 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).

[32] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2015. How do we engage with activity trackers? A longitudinal study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp 2015)*. 1305–1316. https://doi.org/10.1145/2750858.2804290

[33] Steven C Hayes, Kirk D Strosahl, and Kelly G Wilson. 1999. *Acceptance and commitment therapy.* New York: Guilford Press.

[34] Eunkyung Jo, Hyeonseok Bang, Myeonghan Ryu, Eun Jee Sung, Sungmook Leem, and Hwajung Hong. 2020. MAMAS: Supporting parent–child mealtime interactions using automated tracking and speech recognition. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1, Article 66 (5 2020), 32 pages. https://doi.org/10.1145/3392876

[35] Punam Anand Keller, Isaac M Lipkus, and Barbara K Rimer. 2003. Affect, framing, and persuasion. *Journal of Marketing Research* 40, 1 (2003), 54–64. https://doi.org/10.1509/jmkr.40.1.54.19133

[36] Maura A Kenny and J Mark G Williams. 2007. Treatment-resistant depressed patients show a good response to mindfulness-based cognitive therapy. *Behaviour Research and Therapy* 45, 3 (2007), 617–625. https://doi.org/10.1016/j.brat.2006.04.008

[37] Young-Ho Kim, Jae Ho Jeon, Eun Kyoung Choe, Bongshin Lee, KwonHyun Kim, and Jinwook Seo. 2016. TimeAware: Leveraging framing effects to enhance personal productivity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose,

CA, USA) *(CHI 2016)*. 272–283. https://doi.org/10.1145/2858036.2858428

[38] Rafal Kocielnik, Fabrizio Maria Maggi, and Natalia Sidorova. 2013. Enabling self-reflection with LifelogExplorer: Generating simple views from complex data. In *7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. IEEE, 184–191. https://doi.org/10.4108/icst.pervasivehealth.2013.251934

[39] 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* (Amsterdam, the Netherlands) *(MobileHCI 2008)*. 415–418. https://doi.org/10.1145/1409240.1409302

[40] Mark R Leary, Eleanor B Tate, Claire E Adams, Ashley Batts Allen, and Jessica Hancock. 2007. Self-compassion and reactions to unpleasant self-relevant events: The implications of treating oneself kindly. *Journal of Personality and Social Psychology* 92, 5 (2007), 887–904. https://doi.org/10.1037/0022-3514.92.5.887

[41] Andres Ledesma, Mohammed Al-Musawi, and Hannu Nieminen. 2016. Health figures: An open source JavaScript library for health data visualization. *BMC Medical Informatics and Decision Making* 16, 1, Article 38 (2016), 19 pages. https://doi.org/10.1186/s12911-016-0275-6

[42] 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* (Atlanta, Georgia, USA) *(CHI 2010)*. 557–566. https://doi.org/10.1145/1753326.1753409

[43] Ian Li, Anind K Dey, and Jodi Forlizzi. 2012. Using context to reveal factors that affect physical activity. *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 1, Article 7 (3 2012), 21 pages. https://doi.org/10.1145/2147783.2147790

[44] Zilu Liang, Bernd Ploderer, Wanyu Liu, Yukiko Nagata, James Bailey, Lars Kulik, and Yuxuan Li. 2016. SleepExplorer: A visualization tool to make sense of correlations between personal sleep data and contextual factors. *Personal and Ubiquitous Computing* 20, 6 (2016), 985–1000. https://doi.org/10.1007/s00779-016-0960-6

[45] Kai Lukoff, Cissy Yu, Julie Kientz, and Alexis Hiniker. 2018. What makes smartphone use meaningful or meaningless? *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1, Article 22 (3 2018), 26 pages. https://doi.org/10.1145/3191754

[46] Alain Morin. 2017. Toward a glossary of self-related terms. *Frontiers in Psychology* 8, Article 280 (2017), 9 pages. https://doi.org/10.3389/fpsyg.2017.00280

[47] Kristin D Neff. 2003. The development and validation of a scale to measure self-compassion. *Self and Identity* 2, 3 (2003), 223–250. https://doi.org/10.1080/15298860309027

[48] Kristin D Neff. 2003. Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity* 2, 2 (2003), 85–101. https://doi.org/10.1080/15298860309032

[49] Kristin D Neff. 2011. Self-compassion, self-esteem, and well-being. *Social and Personality Psychology Compass* 5, 1 (2011), 1–12. https://doi.org/10.1111/j.1751-9004.2010.00330.x

[50] Kristin D Neff, Kristin L Kirkpatrick, and Stephanie S Rude. 2007. Self-compassion and adaptive psychological functioning. *Journal of Research in Personality* 41, 1 (2007), 139–154. https://doi.org/10.1016/j.jrp.2006.03.004

[51] Kristin D Neff, Stephanie S Rude, and Kristin L Kirkpatrick. 2007. An examination of self-compassion in relation to positive psychological functioning and personality traits. *Journal of Research in Personality* 41, 4 (2007), 908–916. https://doi.org/10.1016/j.jrp.2006.08.002

[52] Kristin D Neff and Roos Vonk. 2009. Self-compassion versus global self-esteem: Two different ways of relating to oneself. *Journal of Personality* 77, 1 (2009), 23–50. https://doi.org/10.1111/j.1467-6494.2008.00537.x

[53] 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* (Oldenburg, Germany) *(MobileHCI 2020)*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3379503.3405662

[54] Jasmin Niess and Paweł W Woźniak. 2018. Supporting meaningful personal fitness: The Tracker Goal Evolution Model. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montréal, QC, Canada) *(CHI 2018)*. Article 171, 12 pages. https://doi.org/10.1145/3173574.3173745

[55] Nanna Notthoff and Laura L Carstensen. 2014. Positive messaging promotes walking in older adults. *Psychology and Aging* 29, 2 (2014), 329–341. https://doi.org/10.1037/a0036748

[56] Gerard Pauley and Susan McPherson. 2010. The experience and meaning of compassion and self-compassion for individuals with depression or anxiety. *Psychology and Psychotherapy: Theory, Research and Practice* 83, 2 (2010), 129–143. https://doi.org/10.1348/147608309X471000

[57] Laura R Pina, Sang-Wha Sien, Teresa Ward, Jason C Yip, Sean A Munson, James Fogarty, and Julie A Kientz. 2017. From personal informatics to family informatics: Understanding family practices around health monitoring. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Portland, OR, USA) *(CSCW 2017)*. 2300–2315. https://doi.org/10.1145/2998181.2998362

[58] Filip Raes, Elizabeth Pommier, Kristin D Neff, and Dinska Van Gucht. 2011. Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical Psychology & Psychotherapy* 18, 3 (2011), 250–255. https://doi.org/10.1002/cpp.702

[59] Amon Rapp and Maurizio Tirassa. 2017. Know thyself: A theory of the self for personal informatics. *Human–Computer Interaction* 32, 5-6 (2017), 335–380. https://doi.org/10.1080/07370024.2017.1285704

[60] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal tracking as lived informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, ON, Canada) *(CHI 2014)*. 1163–1172. https://doi.org/10.1145/2556288.2557039

[61] 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), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

[62] Gh Samaie and Hojjatollah A Farahani. 2011. Self-compassion as a moderator of the relationship between rumination, self-reflection and stress. *Procedia-Social and Behavioral Sciences* 30 (2011), 978–982. https://doi.org/10.1016/j.sbspro.2011.10.190

[63] Jose Luis Santos, Sten Govaerts, Katrien Verbert, and Erik Duval. 2012. Goal-oriented visualizations of activity tracking: A case study with engineering students. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (Vancouver, BC, Canada) *(LAK 2012)*. 143–152. https://doi.org/10.1145/2330601.2330639

[64] 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. https://doi.org/10.1016/j.cpr.2008.10.003

[65] 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* (Montréal, QC, Canada) *(CHI EA 2018)*. Article alt08, 10 pages. https://doi.org/10.1145/3170427.3188401

[66] Lie Ming Tang and Judy Kay. 2017. Harnessing long term physical activity data—How long-term trackers use data and how an adherence-based interface supports new insights. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2, Article 26 (6 2017), 28 pages. https://doi.org/10.1145/3090091

[67] 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), 284–304. https://doi.org/10.1037/0022-3514.76.2.284

[68] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211, 4481 (1981), 453–458. https://doi.org/10.1126/science.7455683