Say You, Say Me

Investigating the Personal insights Generated from One's Own data and Other's data

Yan, Di; Bourgeois, Jacky; Kortuem, Gerd

DOI
10.1145/3679318.3685345

Publication date
2024

Document Version
Final published version

Published in
NordiCHI '24

Say You, Say Me: Investigating Personal Insights Generated from One's Own data and Other's data

Di Yan
d.yan-1@tudelft.nl
TU Delft
Delft, The Netherlands

Jacky Bourgeois
j.bourgeois@tudelft.nl
TU Delft
Delft, The Netherlands

Gerd Kortuem
g.w.kortuem@tudelft.nl
TU Delft
Delft, The Netherlands

Abstract

The design of collaborative personal informatics (PI) has shifted its focus from using one's own data to integrating others' data to enhance self-understanding. In this trend, understanding the effectiveness of the two data sources in facilitating personal insights becomes essential, as a comprehensive understanding of self-understanding requires insights from both individual and interpersonal perspectives. While recent studies have suggested the potential role of others' data as a reflective medium to generate personal insights, little is understood about its distinctive effectiveness in personal insights generated compared to one's own data. To address this gap, we conducted a crowdsourced study involving two participant groups (N1=N2=60) in a data-informed reflection task: Data Providers (DP) reflecting on their own data; Non-Data Providers (NDP) reflecting on the data provided by DP. Analyzing the textual responses, we assess the reflection levels, self-disclosure levels, and characteristics of personal insights. Our findings uncover that others' data possess a comparable effectiveness in facilitating reflection and self-disclosure of personal thoughts and feelings. Others' data displays a strength in supporting value judgments, while one's own data excels in enhancing behavioral awareness. This research sheds light on the design of collaborative PI, offering insights into how to leverage the benefits while mitigating the disadvantages of both data sources to enhance the self-understanding.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; *Empirical studies in collaborative and social computing*.

Keywords

Personal data, Personal Informatics, Self-Reflection

ACM Reference Format:
Di Yan, Jacky Bourgeois, and Gerd Kortuem. 2024. Say You, Say Me: Investigating Personal Insights Generated from One's Own data and Other's data . In *Nordic Conference on Human-Computer Interaction (NordiCHI 2024), October 13–16, 2024, Uppsala, Sweden.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3679318.3685345

NordiCHI 2024, October 13–16, 2024, Uppsala, Sweden

ACM ISBN 979-8-4007-0966-1/24/10
https://doi.org/10.1145/3679318.3685345

1 Introduction

Recent research in HCI has witnessed an increase in the design of collaborative Personal Informatics (PI) systems. In contrast to the common assumption in PI that an individual's knowledge of their personal data facilitates generation of personal insights, recent practices have shifted towards involving others' data to enhance reflection on one's own experiences. This integration of other's data extends to various practices, including incorporating family members' data for providing social and contextual information [58, 66], integrating cohorts' data to compare related behaviors [30, 50, 59], and curating personal data online to stimulate commenting and reflecting on personal experiences [1, 19, 33, 51].

Apart from developing practical approaches for making sense of data, it is important to deepen our understanding of the effectiveness of different data sources in facilitating constructing comprehensive self-knowledge [39, 54]. Different data sources possess various effectiveness for generating personal insights [6, 39], which contribute to the description of different self-images [61]. To achieve a comprehensive self-understanding, it is necessary to consider not only self-images arising from personal experiences spanning the past, present, and future, but also those emanating from interpersonal interactions [60, 61]. Understanding the effectiveness of different data sources for personal insights generation is especially relevant in the context of contemporary PI systems that prioritize comprehensive self-understanding through the integration of others' data [15, 30].

However, prior work in PI has primarily focused on understanding the effectiveness of one's data in facilitating personal insights generation. This has led to a concentration on understanding the reflection process and outcomes related to one's own data [17, 18, 53, 69], such as investigating the reflection levels and insight-gaining patterns through visual exploration of one's data [17] and characterizing types of personal insights derived from making sense of personal data individually [17, 18]. In terms of others' data, recent work has suggested that it can serve as digital representations of individuals for subjective analysis [49] and promote the construction of personal narrative through interpretation and remembrances [24, 32], as opposed to providing an objective truth. Thus, engaging with others' data not only aids in analyzing alternative self-images from an interpersonal perspective [31, 61] but also enriches the self-image by introducing intricate personal narratives constructed through (mis)interpreting and (mis)remembering prompted by data [32, 35]. While those insights suggested the potential role of others' data in the reflective process, an understanding of how insights generated from others' data differ from those generated from one's own data is missing. In this study, we pose the following research question:

RQ: How do the personal insights generated by making sense of and reflecting on others' data differ from those derived from one's own data?

We conducted a study involving a data-informed reflection task wherein participants make sense of sleep data and reflect upon their experiences, via a crowdsourcing platform. We focused on sleep as a context due to the popularity of sleep trackers [22, 43] and prevalent discussions in both online forums [42] and interpersonal Personal Informatics (PI) [58]. We recruited a total of $N = 120$ participants evenly distributed into two groups: Data Providers ($N = 60$) and Non-Data Providers ($N = 60$). Data Providers (DP) comprised participants who were sleep trackers and submitted a screenshot of their sleep data. Non-Data Providers (NDP) were participants who did not submit their sleep data but reported prior experiences in sleep data collection and an interest in understanding their sleep patterns through data. To assess the effectiveness in facilitating personal insights and privacy concerns related to disclosing personal information with data, we systematically evaluated the reflection level, self-disclosure level, and types of insights derived from responses of both DP and NDP.

Our results are threefold. First, others' data possesses comparable efficacy to one's data in facilitating reflective description and dialogic reflection, as well as disclosing personal thoughts, and feelings. Second, others' data demonstrates greater efficacy in assisting individuals in expressing their value judgment by articulating perceptions, attitudes, and opinions. Conversely, one's own data proves more beneficial in enhancing individuals' awareness of their past behavior, particularly in the reconsideration of self-assumptions. Finally, DP tended to generate insights by comparing their self-cognition with data, while NDP tended to compare and interpret other's data to invoke the recall and recognition of their past experiences. These findings underscore the significant role of others' data in fostering reflection, and articulate the distinct effectiveness of others' data versus one's own data in generating personal insights. We provide guidance on leveraging the strengths and mitigating the disadvantages of the two data sources in generating personal insights. We discuss the implications for the design of collaborative PI systems for enhanced insight generation.

2 Related Work

2.1 Collaboration in Personal Informatics

A growing body of literature in HCI explores the design of collaborative Personal Informatics (PI) [20, 41]. One of the notable shifts is from personal health informatics to family health informatics, where the personal data of family members becomes a valuable resource to provide social and contextual information for understanding the interconnected relationship between each family member's behaviors [45, 58, 66]. Furthermore, several studies have embraced the inclusion of others' data as a comparative tool for participants, facilitating the identification of behavioral differences and thereby enriching reflective insights [30, 59]. There is also a rise in online co-curation of personal data, stimulating sharing and reflection on personal experiences [19, 27, 33]. Making sense of and commenting on others' personal data, especially peers who share related experiences, has proven beneficial in enhancing participants' understanding of their own experiences and in promoting the management of well-being and chronic symptoms [1, 19, 33, 51]. In addition, within these collaborative contexts, self-disclosing personal information, thought and feeling upon personal data is a default activity, which benefits a reciprocal process for people to increase self-knowledge and gain emotional support [26]. Prior research has revealed that multiple factors related to data, such as information level [25], post content [26], and presentation style [27], can influence the disclosure of personal data and related experiences.

Apart from exploring the utilization of others' data, it is also crucial to understand the effectiveness of various data sources in facilitating the generation of personal insights [39]. Previous research has emphasized that within a multi-faceted data flow, insights derived from certain data are considered more valuable than others [6, 39]. For example, the more "obvious" insights generated from data, such as "being happier on weekends," provide minimal value in enhancing self-knowledge [6]. Especially for constructing comprehensive self-understanding, it becomes essential to derive personal insights from data that describe multi-faced self-images, including those arise from personal experiences spanning past, present and future, as well as from interpersonal interactions[61]. Thus, it is important to have a deeper understanding of the different effectiveness of others' data and one's data in facilitating personal insights, especially for contemporary PIsystems that prioritize comprehensive self-understanding through the integration of others' data.

2.2 Sensemaking and Reflection on Personal Data

In the realm of personal visualization and Personal Informatics, making sense of and reflecting on personal data a pivotal activity that empowers individuals to gain personal insights, thus increasing self-knowledge and potentially enacting behavioral changes [15, 17, 18, 47]. In personal visualization, making sense of personal data through visual exploration is the essential approach to facilitate self-reflection [17, 18]. In line with that,Li et al. [47] proposed a model where reflection is an integral part of a comprehensive five-stage process (Preparation, Collection, Integration, *Reflection*, and Action) where individuals make sense of personal data visualizations to generate insights. While the later models emphasized that reflection happens associated with the lived experiences, making sense of personal data (and visualization) remains a key activity to provide reflective material [28].

Numerous studies in personal visualization and personal informatics have investigated the value of one's own personal data in supporting reflection [15, 21–23]. Within this domain, researchers have delved into understanding the reflective process and its outcomes. For instance, Choe et al. [17] applied a taxonomy of five reflection levels to investigate how visual exploration on one's own data support reflection. Their study revealed that visual exploration on one's own personal data predominantly facilitates *description* (R0) and *descriptive reflection* - the two low levels of reflection refer to revisiting of past experiences and revisiting with an explanation of past experiences, respectively. Furthermore, they also observed emergence of *dialogical reflection* (R2), which refers to the exploration of relationships among ideas and experiences with the aim of deriving generalizations and attaining. The *transformative reflection* (R3)-characterized by questioning initial self-assumptions and shifts in fundamental self-understanding or behavioral practices and the *critical reflection* (R4) referring to reflecting on aspects that transcend the immediate context (e.g., social and ethical issues) appeared to be rare. In addition, their findings underscore that data serves a dual role in reflection process: recalling past behaviors as well as external contexts and prompting new questions for further exploration. Furthermore, a few studies investigate the type of personal insights. For instance, Choe et al. [18] proposed a framework to examine the characteristics of personal insight as an outcome of self-reflection.

In this study, we extend the exploration of reflective practices to encompass the reflection on others' data. We adopted the reflection taxonomy and characteristics of personal insights from previous works [17, 18] when considering the effects of reflecting upon others' data against one's own data.

2.3 The Authority of Data

Personal informatics and other widespread uses of data in society have prompted research inquiries into the authority of data, questioning the unique power of insights derived from data (and visualization) and how they complement other types of insights. One perspective, rooted in the concept of "data doubles" [65], posits that personal data serves as a digital representation of individuals, which is not a source of objective truth but amenable to figurative reconstruction for subjective purposes like personal reflection and interaction. Any attempt to define a meaning from data involves the performance of agential cut, where people separate data into elements or characters from a dataset according to their subjective conceptual boundaries [49]. Furthermore, Rapp and Tirassa [61] suggested that involving alternative perspectives, such as presenting personal data of others, can boost the empathy toward one's own experiences. This, in turn, contributes to the creation of interconnected self-images of oneself that come from the social interactions, as one of the four facets of thyself. We reference these perspectives to suggest that alternative sources of personal data, such as that of others, can be considered as material for constructing self-identity.

Moving beyond exploring the analytic value of personal data, recent studies adopt a different perspective, viewing it as a creative and communicative artifact that facilitates the construction and elicitation of personal narratives through (mis)interpretations. For instance, Gulotta et al. [35]'s "Curatorial Agents" reveals differences between human-created and machine-created interpretations of data, and points out that these differences can be productive and that misinterpretations (and mis-remembrances) are important phases in how people create and share the narratives of their lives. In the "Metadating" project, researchers organized a speed dating event where participants crafted and exchanged data profiles containing various types of data, including entirely accurate data collected by tracking devices, estimated data, and fabricated data [24]. It illustrates the concept that data can function as a "creative material" with its unique "social life," departing from its conventional role as a source of objective truth. Similarly, Friske et al. [32]'s project involves two participants creating and interpreting each other's knitted data representation, revealing that developing one true narrative towards multiple narratives from data equally informs both participants in understanding the data.

The above insights suggest that others' data has the potential to facilitate reflection, serving as a medium to provide resources for subjective analysis of past behavior and promoting the construction of personal data narratives through (mis)interpretation. Despite these findings, little is known about the distinctive effectiveness of others' data in generating personal insights. Our work sets out to address this gap, by delving into a comparative study on understanding the differences in the reflection levels, self-disclosure levels, and characteristics of personal insights derived from one's own data and others' data.

3 Study

In this study, we aim to understand the different effectiveness of one's own data and others' data in facilitating personal insights generation. We conducted a between-subject study using the crowdsourcing platform Prolific, recruiting two groups of crowd-workers ($N = 120$, $N_{DP} = 60$, $N_{NDP} = 60$). The Data providers (DP) comprised sleep tracker participants who submitted a screenshot of their sleep data, while the Non-Data Providers (NDP) did not provide data, but had experiences in sleep data collection and expressed an interest in understanding their sleep through data. Both groups engaged in a data-informed reflection task (see section 3.4) where they provided reflective responses in textual format by making sense of their own data and others' data respectively. Employing a combination of quantitative and qualitative methods, we systematically characterized the mechanisms and outcomes of sensemaking and reflection on both one's own data and others' data.

3.1 Study Context

We conducted the study in the context of sleep for two main reasons. Firstly, sleep is a dual-natured activity—partially a bodily process beyond conscious control, and partially a self-conscious activity influenced by daily routines and social family activities [14, 42]. This dichotomy has sparked extensive discussions within online communities [42] and interpersonal PI [58], wherein people share their sleep experiences, offer collaborative support, and gain a deeper understanding of bodily issues. Secondly, sleep data is a common data type that describes direct and simple human behaviors (e.g., asleep, awake) throughout the sleep process [43, 62], which is easy to understand for both DP and NDP. Thus, it can serve as equitable material for both DP and NDP to reflect upon their experiences.

We collected screenshots of Apple Health sleep data[1] from DP as reflection material for both DP and NDP (see the example in Figure 1 on the left-hand side). According to the informatics design guidance for reflection [21], the Apple sleep data report is considered good reflection material for the following reasons. *1*) It uses bar charts to represent sleep time with simple statistics (e.g., sleep goal, average sleep time) of the past week, which is easy for both types of participants to understand even at a glance. *2*) The bar chart of sleep time in the recent week is given in time series, which can help participants notice and reflect on changes in their sleep time that they might not perceive otherwise. *3*) It includes all types of

[1]Apple Health: https://www.apple.com/ios/health/

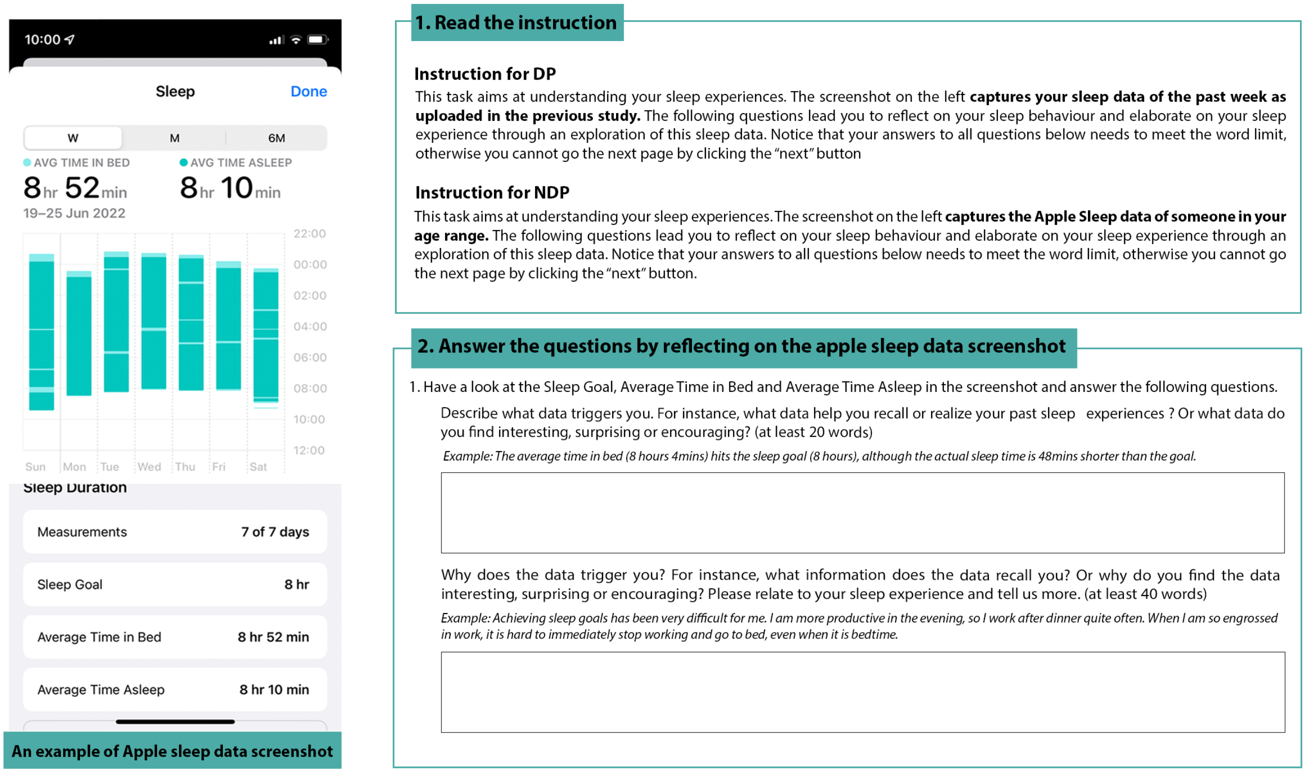


Figure 1: Illustration of the data-informed reflection task for DP and NDP. The instructions on the top right slightly differ (bold font) for DP and NDP, while the screenshot and questions are the same.

anchors to support various reflection levels such as average values (e.g., average sleep time), extreme values (e.g., latest sleep time), and patterns and trends (e.g., change of sleep time in one week). *4*) The Apple sleep data is sufficiently detailed yet rich enough, as it does not record behaviors and experiences about which individuals may have limited knowledge or may even be unaware of themselves. For instance, other apps recording unconscious sleep behaviors (e.g., sleep phases) are overly complex to make sense of, thus not efficient in facilitating self-reflection.

3.2 Recruitment and Participants

We recruited 120 crowd workers ($N = 120$) through the crowdsourcing platform Prolific[2], comprising 60 DP and 60 NDP. DP participants were crowd workers who consistently tracked their sleep using a wearable device every day in the week preceding the study. Recognizing the diverse motivations behind (sleep) data collection [43, 56, 62], we welcomed crowd workers who had maintained regular sleep data collection, rather than imposing constraints for specific data collection purposes. NDP participants did not provide their data, and they were individuals who reported prior experience in collecting sleep data and interest in understanding their sleep patterns through data. To further foster reflection, we categorized DP and NDP into three 10-year age brackets (20-29, 30-39, 40-49).

The recruitment criteria and grouping strategy for DP and NDP were chosen to balance the complexity of providing meaningful data for reflection and executing the experiment. Previous literature has emphasized the importance of meaningfulness, which refers to people's interest in data and its relatedness to their lives, rather than merely seeking familiarity between participants [44]. While prior research often involves individuals with close relationships (e.g., family members and colleagues) to encourage reflection [42, 60], given the challenge of recruiting participants in close relationships, we argue that our selection of NDP is also valid. Self-trackers who are not acquainted but share common interests and self-tracking behaviors are considered to share a strong sense of relatedness in PI research [1, 19, 37]. Individuals from the same group can offer alternative perspectives in relating personal experiences and even breaking social norms [1, 61]. As for the grouping strategy, we prioritized the key factor - age - that affects sleep behavior [29, 48, 63], considering the challenge of controlling multiple influencing factors (e.g., occupation, gender, and health condition). We specifically grouped participants into 10-year age ranges, as prior sleep literature has considered 10 years a representative period for investigating the influence of age on sleep quality [48, 52]. Thus, grouping participants by 10-year age ranges promotes the establishment of connections and facilitates reflection.

[2]Prolific crowdsourcing platform: https://prolific.co

To ensure the recruitment criteria, we utilized two open-ended questions for NDP to elicit explanations regarding 1) their past experiences with data collection and 2) their motivations for reflecting on their sleep patterns. Similarly, we employed an open-ended question to inquire about DP's data collection experiences and objectives. After coding the responses to these questions, we observed that among NDP, 45.3% had previously collected sleep data, while 54.7% reported occasionally or regularly collecting sleep data presently. The main motivations for data collection, identified for both DP and NDP, included sleep management, daily activity management, evening baby care-giving management, and illness management.

To ensure response quality and mitigate spam, participation was restricted to workers with a minimum acceptance rate of 95%. Moreover, we exclusively recruited participants who were native English speakers to ensure that NDP could comprehend the screenshots collected from the DP. Participants were compensated at a rate of 8.00 pounds per hour, which was deemed competitive according to the platform's standards. Our institution's Human Research Ethics Committee and Privacy Team conducted a thorough review and approved these activities.

3.3 Procedure

We divided the study into two phases.

- *Phase 1a – Data uploading task, 3 minutes/* ***Data Provider***. We recruited DP and assigned them to capture and upload a screenshot of their Apple sleep data. To guide DP through the process, we provided detailed step-by-step instructions, emphasizing the avoidance of personal identifiers in the screenshots. As an incentive to enhance the likelihood of obtaining valid screenshots, a bonus of 0.5 pounds was offered. This phase produced over 80 screenshots from DP, including over 20 for each of the three age ranges. The first author manually verified the validity of each data screenshot.
- *Phase 1b – Data-informed reflection task, 15 minutes/* ***Data Provider***. We extended invitations to DP who had successfully provided valid screenshots in the previous phase 1a. They were prompted to engage in a 15-minute data-informed reflection task (see details in Section 3.4). First, they were asked to explain their motivation for collecting sleep data by answering an open-ended question. Then, they were prompted to make sense of their Apple sleep data (screenshot provided in Phase 1a) and disclose their sleep experiences by answering the open-ended reflective questions. The response rate was generally high, with only a small fraction of DPs not responding to our invitation. Upon collecting 60 responses from DPs, we concluded this task. This phase yielded a total of 60 responses from DP, consisting of 20 responses in each of the three age ranges.
- *Phase 2 – Data-informed reflection task, 15 minutes/* ***Non-Data Provider***. In this phase, we called for NDP to execute the data-informed reflection task (see details in Section 3.4). To ensure that NDPs met the study requirements, we first asked them to report their past experiences in collecting sleep data and their motivation for reflecting on their sleep via two open-ended questions. Subsequently, presented with a randomly selected data screenshot from a DP (Phase 1) within the same age range, NDP were encouraged to utilize this screenshot to reflect on and share their own past sleep experiences by responding to the reflective questions. The quality of responses was assessed manually by the first author. Three responses were excluded from participants who reported no prior data collection experiences and lacked interest in understanding sleep, while two responses were discarded due to low-quality input in the reflective questions (primarily consisting of short answers with only 5 words in all questions). After recruiting 5 more participants to execute the task, this phase yielded 60 responses from NDP.

Participants completed all tasks via a web platform we developed in Python Django and hosted on [Author's University]'s servers. Before each task, we provided participants with a consent form describing the task in detail, the expected completion time, and the appropriate data-sharing policies. For DP, we specifically included a clause for sharing their anonymized data screenshot with other crowd workers. Finally, we scrutinized data submissions at each phase to ensure the anonymity of provided screenshots and written text.

3.4 Data-informed Reflection Task Design

The data-informed reflection task involves participants making sense of and reflecting upon sleep data by responding to open-ended reflective questions (see Figure 2). Recognizing that establishing a rationale for reflection is crucial for directing reflection toward the intended outcome [31], distinct instructions with two different purposes were provided to DP and NDP participants, accompanied by a screenshot of sleep data (see top right of Figure 1). DP were instructed to reflect on their *own* sleep data and disclose their *own* sleep experiences. In contrast, NDP were prompted to make sense of *someone else's* sleep data but disclose and reflect on their *own* sleep experiences.

Both DP and NDP were presented with the same open-ended questions (Figure 2) to reflect on and disclose their feelings and experiences of sleep. We carefully designed these reflective questions by striking a balance between providing guidance and allowing freedom to facilitate reflection. This decision is informed by prior literature emphasizing that reflection often requires explicit guidance rather than occurring naturally [7, 31].

First, we chose open-ended questions to support data sensemaking and reflection. Open-ended questions provide a standard and flexible way to explicitly guide and structure reflection [21, 31]. They prompt participants to specifically consider issues relevant to achieving the intended purpose of reflection while allowing for dynamic levels of reflection. Furthermore, we intentionally divided the reflection questions into sub-questions to guide participants from lower levels to higher levels of reflection. Specifically, the first sub-questions are designed for participants to identify data patterns from the visualization, responding to reflection levels from lower (R0, R1). The second sub-questions are designed to prompt people to relate the identified data patterns with their personal experiences, responding to higher reflection levels (R2, R3). This decision is informed by prior literature highlighting higher levels of reflection are usually supported and prepared by the lower levels of reflection [31]. In addition, we avoided incorporating reflection level R4

A. Have a look at the Sleep Goal, Average Time in Bed and Average Time Asleep in the screenshot and answer the following questions.

- *Describe what data triggers you. For instance, what data helps you recall or realize your past sleep experiences? Or what data do you find interesting, surprising or encouraging? (at least 20 words)*
- *Why does the data trigger you? For instance, what information does the data recall you? Or why do you find the data interesting, surprising or encouraging? Please relate to your sleep experience and tell us more. (at least 40 words)*

B. Have a look at the sleep trend and pattern in the bar chart and answer the following questions.

- *Describe what data triggers you. For instance, what data helps you recall or realize your past sleep experiences? Or what data do you find interesting, surprising or encouraging? (at least 20 words)*
- *Why does the data trigger you? For instance, what information does the data recall you? Or why do you find the data interesting, surprising or encouraging? Please relate to your sleep experience and tell us more. (at least 40 words)*

C. Have a look at the latest sleep time and earliest wake-up time in the bar chart and answer the following questions.

- *Over the recent month, do you usually sleep at the latest sleep time or earliest wake-up time in the screenshot? If you did, what were the reasons for going late to bed or waking up early on that day(s)? If you didn't, please tell us why you do not sleep or wake up at these times? Please mention yes or no at the beginning of your answer(at least 40 words).*

Figure 2: The reflective question list for DP and NDP

(the highest), as it tends to be excessively abstract and detached from real-life scenarios [31]. Figure 2 provides detailed information on our list of three reflective questions in the data-informed reflection task.

- **Question A:** This question started by prompting users to identify patterns in their overall sleep data (e.g., sleep goal, average time in bed, and average time asleep), providing an overview of their sleep experiences. The first sub-question aims to guide participants in identifying and explaining triggering data points that capture their behaviors, corresponding to description(R0) and descriptive reflection(R1). Sample questions are provided to further specify our intention. The second sub-question is designed to elicit the reasons behind participants' identified or recalled behaviors. By encompassing the overall data, this question investigates people's goals and plans, which can encourage reflection not only on the relationship between experiences but also on behavioral change [67], responding to the levels of dialogic reflection (R2) and transformative reflection (R3). Sample questions are also provided to give further detailed instructions.
- **Question B:** This question began by inquiring about detailed sleep data (e.g., sleep trends and patterns) that represented a more concrete picture of sleep behaviors. Similar to Question A, the first sub-question was designed to guide participants in identifying and explaining data patterns that drew their attention, while the second sub-question prompted participants to provide reasons behind the identified or recalled behaviors. Considering the use of sleep data patterns and trends as anchors, the second sub-question can encourage participants to reason about the relationships within multiple experiences [6], responding to dialogic reflection (R2).
- **Question C:** This question utilized outlier data (e.g., latest sleep time and earliest sleep time) as an anchor to guide reflection. The first sub-question ("Over the recent month, do you usually sleep at the latest sleep time or earliest wake-up time in the screenshot?") prompted participants to identify extreme data points and recall their past experiences beyond the presented data, addressing description (R0) and descriptive reflection (R1). Inquiring about the reasons behind these extreme behaviors, the follow-up questions encouraged participants to delve into detailed context and information in interpreting and relating the data [6], fostering dialogic Reflection (R2).

3.5 Data Analysis

We collected textual responses from the data-informed reflections on sleep provided by DP and NDP. By employing a combination of quantitative and qualitative methods, we analyze reflection levels, self-disclosure levels to understand the effectiveness of two data sources as reflective medium. In addition, we analyzed the self-disclosure level to gain insights into the influence of different data sources on people's disclosure of personal information, thoughts, and feelings.

In the first round of analysis, we aimed to gain an overview of the difference in reflection and self-disclosure levels between DP and NDP. The first two authors coded the collected annotation by applying the frameworks of Fleck and Fitzpatrick [31] and Barak and Gluck-Ofri [4] to identify reflection and self-disclosure levels. We specifically applied the self-disclosure framework proposed by Barak and Gluck-Ofri [4], as it is designed and widely adopted for analyzing self-disclosure in the online environment. For each participant, we rated the reflection and self-disclosure level three times, corresponding to the answers to the three questions. The first two authors coded 30% of the reflections separately and then resolved disagreements through discussions. After agreeing on the first 30% of the codes, they independently coded the remaining reflections, compared their codes, and discussed possible revisions. This process resulted in final inter-rater reliability of 98%. Notice that we did not evaluate R0, due to the fact that R0 involves the recall of past experiences without further explanation, a subtlety that can be challenging to observe and distinguish from the text responses. We also excluded the evaluation of R4, as it refer to the reflection on social and ethical aspects that is rare. Table 1 and 2 provides examples quotes for reflection and self-disclosure levels. Finally, we compared the distribution of reflection and disclosure

Table 1: Example Quotes Categorized by Reflection Level

Reflection Level	Example Quote
Reflective Description (R1)	"No, I usually sleep very late. Mainly past midnight. I find myself very busy during the day with my children, housework, and jobs. By the time I got chilled time to myself, it is quite late already." NDP22(30-39)
Dialogic Reflection (R2)	"My bedtime is inconsistent and a lot of nights I have broken sleep. I carry out most of my work in the evenings and I need some time to wind down before bed. I also have a young child who doesn't always sleep well. I need to be up by 7am to get him ready for school." DP3(40-49)
Transformative Reflection (R3)	"I'm in bed for an average of 8 hours 42 minutes. I think I should not be on my phone before I go to bed and wake up at 9. Or I need to change my alarm to wake me up at 9." DP7(30-39)

Table 2: Example Quotes Categorized by Self-Disclosure Level

Self-Disclosure Level	Information	Thoughts	Feelings
No self-disclosure	"Within a short time, the person falls asleep and then wakes up. They get out of bed straightforwardly." NDP60(40-49)	"This person got less sleep on Sunday, potentially because they do not have a strict schedule on weekdays." NDP11(20-29)	"I keep a consistent rhythm with sleep. The graph really shows that. I don't know what else to say." DP37(30-39)
Little self-disclosure	"I think the difference between my time in bed and my sleep time is a lot bigger than this data shows." NDP29(30-39)	"I know I sleep badly. This data just proves it." DP24(30-39)	"It is frustrating to see tangibly the effect of commuting and being in the office on my sleep and life." DP35(30-39)
High self-disclosure	"I take medicine at night, and I have to eat something before taking it. This can sometimes push my bedtime too late." DP13(30-39)	"I think I need to adjust my bedtime to 11 p.m., or maybe stay active until around 10 p.m." DP33(30-39)	"I wake up frequently during the night because I have a young child waking throughout the night. It is depressing to see the actual gaps in my sleep." DP31(30-39)

levels of the two participant groups, and calculated the p-values using a Mann-Whitney-U test.

In the second round of analysis, we combined inductive and deductive coding based on the data-driven personal insights categories [16, 17], to understand the types of personal insights participants gained from data-informed reflection. For example, we extracted the following piece of annotation from DP2:"I have been sleeping a little less than the recommended 8 hours and it mentally shows. I would like to get at least 8 hours and not." This quote involves two types of insights: *confirmation* ("...it mentally shows...") and *against external data* ("less than the recommended 8 hours"). The two first authors separately coded the first 30% of the reflections. Then, they discussed discrepancies, revised and expanded the existing categories until they reached an agreement of 90%, thereby creating a new coding scheme of personal insight categories. After this step, the first author applied the coding theme to the rest of the annotations, generated the final coding theme of personal insight categories after several iterations.

4 Results

This section presents our results, mapping differences and similitudes between Data Providers (DP) who reflect on their data from Non-Data Providers (NDP) who reflect on others' data.

4.1 Reflection Levels

We observed no significant difference in the average reflection levels between the answers from the two participant groups (Mann-Whitney $U = 15945$, $N_{DP} = N_{NDP} = 60$, $p\text{-}value > .05$, two-tailed). To further compare the distribution of reflection levels between the two groups, Table 3 shows the proportion of answers provided by DP and NDP at each reflection level. We observed that the answers of DP are relatively balanced across all three reflection levels, with more answers given at the R1 level than those at the R2 and R3 levels. On the other hand, NDP answers show a more skewed distribution: answers at the R2 level ($NDP_{R2} = 60\%$) are significantly more than the others, very few answers are found at the R3 level ($NDP_{R3} = 8.33\%$).

These results indicate that *1*) DP can reflect at all three levels, describing and explaining past behaviors and experiences, reconsidering self-assumptions, and new insights. In contrast, NDP can only reflect at the levels of R1 and R2, lacking consideration of personal assumption and intention to change behavior (R3). *1*) DP provide more direct descriptions and explanations of experience without exploring alternate explanations (R1), whereas NDP tend to explain the relationships between experiences and other points of view (R2).

4.2 Disclosure Level

We now compare the types (information, thoughts and feelings) and the degree (no, little or high) of disclosure of answers by DP and NDP. Overall, DP shows a significantly higher degree of information

	Reflective Description (R1)	Dialogic Reflection (R2)	Transformative Reflection (R3)
DP	44.44%	27.78%	27.78%
NDP	31.67%	60.00%	8.33%

Table 3: Distribution of reflection levels

	Information			Thoughts			Feelings		
	No	Little	High	No	Little	High	No	Little	High
DP	0.00%	17.28%	82.72%	44.44%	29.01%	26.54%	67.90%	30.86%	1.23%
NDP	2.47%	28.40%	69.14%	43.21%	30.86%	25.93%	75.31%	24.07%	0.62%

Table 4: Distribution of disclosure degrees of information, thoughts and feelings

disclosure than NDP (Mann-Whitney U = 13962.5, $N_{DP} = N_{NDP}$ = 60, *p-value* < .05, two-tailed), but no significant difference in the degrees of thoughts (Mann-Whitney U = 16044, $N_{DP} = N_{NDP}$ = 60, *p-value* > .05, two-tailed) or feelings (Mann-Whitney U = 15282.5, $N_{DP} = N_{NDP}$ = 60, *p-value* > .05, two-tailed) disclosure. Table 4 shows the degree distribution of answers given by DP and NDP across all three disclosure types. We observed similar degree distribution of DP and NDP answers for all three disclosure types, except for a slightly higher degree of DP in disclosing information.

These results suggest that DP and NDP have a comparable tendency to disclose personal thoughts and feelings. The only difference is that DP has a slightly stronger tendency to provide detailed personal information such as past behaviors and personal context.

4.3 Different Insight Types

To classify insights by types in the directed content analysis, we enriched the classification of Choe et al. [17] to include the different insight types for DP and NDP, summarized in Table 5. We start by enriching the insight type of *recall*, which covers insights generated by recalling past behaviors and events not captured by data. Next, we refined the previous sub-insight type *external data* from Choe et al. [17] with **behavioral description**, **reasoning**, and **life events** and **justification with external data** which both DP and NDP share. Then specifically to NDP, we identified the sub-types **difference** and **similitude** while reserving the original *confirmation* and *contradiction* of Choe et al. [17] for DP. We refined the *value judgment* type by distinguishing between the judgment over subjects, namely **data judgment** and **behavioral judgment**. Finally, we added **Behavioral awareness** as a new insight type for both DP and NDP.

With this extended classification, we identified close numbers of insights for both DP (720 insights, average=12/person) and NDP (600 insights, average=10/person). We characterize nuances in the following.

4.3.1 Different composition of recall. *Recall* is the most frequent insight type and its amount is similar for both DP and NDP. DP generated 242 recall insights while NDP generated 253 insights, accounting respectively for 37.6% and 43.5% of the total amount of insights. Despite this similar amount of recall insights, sub-insights compositions differ between DP and NDP. Only DP generate the sub-insight of *life events*, *contradiction* and *confirmation*, and they also create more sub-insights of *behavioral description* (105 insights for DP, 80 insights for NDP). NDP generated a fair amount of sub-insight *difference* (87 insights for NDP) and *similitude* (50 insights for NDP). As shown below, NDP7 explain their early wake-up time by comparing it to the late sleep time of the DP.

Difference: "*I find the data surprising that someone is going to bed so late and waking up late in the day. I have to wake at 6 am every day to get ready for the day and do household chores.*" NDP7 (20-39)

Similitude: "*The average time in bed and average time asleep triggers me as I am very similar in that it will take me a while to fall asleep, despite actually being in bed. However, it is encouraging to know I am not alone in this.*" NDP17 (20-29)

It indicates that both DP and NDP can use data as an anchor to recall their past behaviors. DP recall their memory of past events, behavior, and experiences directly triggered by data. In contrast, NDP recall their past behaviors by comparing with the behavior represented by data through the identification of differences and similitude.

4.3.2 NDP express stronger judgment. We identified two distinct sub-insights under Value Judgment: *behavioral judgment* and *data judgment*. These two sub-insights delineate between the subjects of judgments, with *Behavioral judgment* conveying positive and negative perceptions, attitudes or opinions about the behavior represented by data, and *data judgment* expressing a perception on the value of the data itself. Our analysis reveals that DP tended to comment on their own data and share opinions directly.

Data judgment: "*Analyzing the data is always helpful when you look at it in the big picture. The data helps to understand sleep patterns and helps to show how I can improve my sleep.*" DP47 (40-49)

NDP express their own perceptions, understanding and opinions on sleep behavior by commenting on other's data or the behavior represented by data.

Behavioral judgment: "*I find the data surprising that someone is going to bed so late and waking up late in the day. I have to wake at 6am every day to get ready for the day and do household chores. I also find it surprising that the person falls asleep so quickly after getting into bed.*" DP12(40-49)

Figure 3a highlights that, overall, NDP generate substantially more value judgment insights than DP. In particular, DP generate twice as many behavioral judgment insights as NDP (47 insights for

Table 5: Personal insight types for DP and NDP. Note that '(*DP*)' represent the insights only generated by data providers, and '(*NDP*)' represent the insights only generated by non-data providers

Type	Subtype	Description	Example quotes
Recall	Behavioral description	Remind and describe past behavior not captured by data but reminded by data	"I notice that I have a hard time staying asleep all night. I am often waking up and rolling over back and forth during the night." DP31 (30-39)
	Reasoning	Explain the reasons behind the behaviors, such as habits, life condition, routines...	"I usually always wake up at the same time every day because I work every day and have to get up at that time." DP9 (20-29)
	Difference (NDP)	Describe own behavior contrasting with the observed data	"The similar bedtime on the weekends surprises me as I often go to sleep several hours later on those days compared to the weekend." NDP2(20-29)
	Similitude (NDP)	Describe own behavior similar with or related to observed data	"It appears as if towards the weekend, the time that they went to sleep was later which is similar to my sleep routine as I have to get up earlier on the weekdays." NDP11 (20-29)
	Life events	Explain data points by recalling and elaborating on past event	"For the late sleep time on the Saturday I was very restless and could not get to sleep, thinking back I had a few drinks with caffeine in them which may have been a factor." DP43 (40-49)
	Justification with external data	Bring external data to justify behavior or opinions	"I have my alarm at 8:55 but will always snooze till 9 AM or a little past 9 AM to start my day." DP35 (30-39)
	Confirmation (DP)	Collected data confirms existing knowledge	" This isn't surprising as the weather has been very warm which has had a negative impact on my sleep." DP11 (20-29)
	Contradiction (DP)	Collected data contradicts existing knowledge	"I am fairly diligent in terms of going to bed at a similar time, the varying state of the time that I rise is surprising to me." NDP2 (20-29)
Value Judgment	Data judgment	Convey positive or negative connotations about the measured data	" I feel like this information is helpful in knowing how well I slept." DP4 (20-29)
	Behavioral judgment	Convey positive or negative connotation about the behavior represented by the data	"5 am seems a horribly early time to wake up!" NDP23 (30-39)
Behavioral awareness	Behavioral wish	Describing behavior that the person would like to have	"I really wish I could have more consistent sleep but sometimes I stay up too late to hang out with friends." DP5 (20-29)
	Intention to change	Express the intention and the motivation to change behavior	"As it shows how bad my sleep schedule is. It makes me want to change my habits." DP9 (20-29)
	Changed self-assumption	Adjust or change the understanding of themselves	"I was also surprised by the fact that I am waking up in the night. I thought I consistently slept through the night with no problems." DP28 (30-39)

DP, 89 insights for NDP), and the amount of data judgment insight is similar for both crowds. It indicates that NDP have a stronger ability to express their perception, attitudes and opinions by reflecting on others' data. It also stresses the critical role of others' data in helping NDP disclose their perception of bad, average, and good behaviors.

4.3.3 DP gain more behavioral awareness from their data. We augmented the insight types from Choe et al. [17] with *Behavioral awareness* insights, which include three sub-insights: *behavioral wish*, *changed self-assumption* and *intention to change*. Figure 3b highlights that DP gain more insights into behavioral awareness than NDP (70 insights for DP, 30 insights for NDP). Specifically, DP discloses almost 3 times more *change self-assumption* insights than NDP (45 insights for DP, 12 insights for NDP), while the number of *intention to change* insights are similar (12 insights for DP, 10 insights for NDP).

Changed self-assumption: "*I'm fairly pleasantly surprised by my averages as I thought they would be much worse, this hasn't been my greatest week for constant sleep.*" DP3 (20-29)

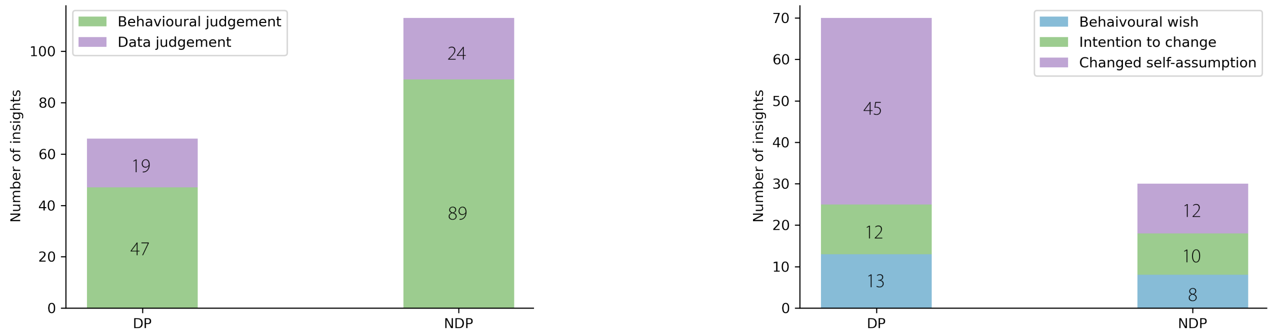
These results indicate that data support DP gain more behavioral awareness, and it exerts influence on DP primarily by making them re-considerate their self-assumption rather than directly initiate behavior change.

4.4 Additional Insights from NDP

Except for the extended classification, we identified two types of additional insights generated only by NDP: **behavioral identification** and **speculation**. Behavioral identification involves NDP identifying purely behavior of DP from their data, without any personal information, thoughts, or feelings of DP. These insights usually prepare a context for NDP to relate to their experiences, opinions and attitudes that are different from or similar to DP.

The second type of insight is speculation, containing speculation and interpretation of DP's behaviors. Through these speculations, we can still discern the lifestyle or past behaviors of the NDP. For example, in the following quote, NDP5 speculates that the time the DP spent in bed before falling asleep was on the phone, implying that they might have the habit of using a phone in bed themselves.

These two additional insights suggest that NDP convey their experiences and behaviors by relating to and interpreting others'



(a) The number of sub-insights of Value Judgment from DP and NDP. This stack chart shows that NDP gains more insights into behavioral judgment than DP, but the number of data judgment are similar.

(b) The number of sub-insights of Behavioral awareness from DP and NDP. This stack chart shows that DP gains more insight into behavioral awareness than NDP, with a big advantage in the number of changed self-assumption.

Figure 3: Comparison of DP and NDP in value judgment and behavioral awareness.

data, diverging from the comparison mechanism outlined in Section 4.3.1.

5 Discussion

In this section, we discuss the differences and similarities between reflection levels, self-disclosure levels and types of personal insights derived from others' data and one's own data. We also discuss the implications of these insights for the design of collaborative personal informatics systems.

5.1 Comparable Value in Facilitating Reflection

In Section 4.1, our findings reveal that there is no significant difference in average reflection levels between DP and NDP. DP participants tended to reflect across all three levels, while NDP participants primarily engaged in reflective description (R1) (60%) and dialogic reflection (R2) (31.67%). This suggests that others' personal data holds comparable value to one's own data in facilitating self-reflection. Specifically, others' data proves valuable in aiding individuals in recalling and explaining experiences (reflective description R1) and in reasoning about the relationship between underlying expectations, needs, and feelings (dialogic reflection R2). Furthermore, in Section 4.3, the close numbers of insights for DP (720 insights, average=12/person) and NDP (600 insights, average=10/person) also suggest echoing this finding of the comparative ability of others' data to one's own data in facilitating insights generation.

Our findings indicate that other's data can serve as a protagonistic material for people to recall and reflect upon experiences, even in the absence of one's own data. This insight contrasts with numerous existing PI tools that focus on the use of one's own data as reflective material [15, 57, 61]. Traditionally, the process of data collection has been considered a crucial step in PI tools, often posing barriers to reflection [8]. Future research could explore innovative ways to incorporate others' data into PI tools, thereby extending the scope of reflection to a broader range of individuals who may not engage in self-tracking. For example, future collaborative PI tools could involve the collection of a small dataset and subsequently share this data with users who share similar interests and experiences within a group setting (e.g., a workplace or educational institution [50]). Moreover, the inclusion of others' data can enrich the reflective process by providing users with a wider range of perspectives and experiences to draw upon. Machine-assisted reflection systems, which excel in providing explanations and interpretations of personal data [40, 64], could leverage others' data as input to offer users a detailed portrayal of different lifestyles and contexts. By providing users with a comprehensive understanding of others' lives, these systems can facilitate deeper reflection and insight generation.

Beyond the realm of PI, our findings also suggest the potential for scaling out data work, such as articulation work [69] and data-enabled design methods [10], by leveraging others' data as material for users to reflect and disclose their past experiences. Involving product users in making sense of personal data has been recognized as a key activity in these fields to reveal underlying expectations, feelings, and experiences [13, 34, 69]. However, current methods often entail significant design and setup efforts for data collection and participant recruitment, limiting their scalability and reach [11]. Our findings suggest that future work in articulation work and data-enabled design could leverage a smaller number of pre-collected data to engage a larger number of participants, thus enabling a deeper understanding of users on a broader scale.

5.2 Differences in Behavioral Awareness and Value Judgment

Our findings in Section 4.3.3 reveal the proficiency of DP in generating insights into "Behavioral awareness," while NDP exhibit a tendency to generate insights related to "Value judgment." These results suggest the different strengths of one's own data and others' data in facilitating insight generation:

One's own data holds greater strength in aiding individuals to enhance their awareness of past behaviors. This insight aligns with a substantial body of prior research in Personal Informatics (PI), rooted in the ego-centric perspective derived from behavioral change theories, which consistently emphasize the pivotal role of one's own personal data in facilitating reflective processes and promoting self-awareness [15, 61]. Therefore, our findings confirm prior PI design principles and underscore the importance of prioritizing the inclusion of individuals' own data when designing PI tools aimed at fostering self-awareness and behavioral changes. PI tools that focus on self-experiments [3, 22] and machine-assisted reflection [40] should prioritize the use of personal data to facilitate the (re)construction of detailed self-images, rather than relying on others' data, which primarily serves to define norms.

Others' data is proficient in assisting individuals in expressing and justifying their perceptions, attitudes, and opinions on behaviors. This finding extends prior research emphasizing the reference value of others' data in establishing norms [30]. Combined with the finding that NDP is comparable in disclosing thoughts and feelings except for more personal information (Section 4.2), others' data is indicated to be valuable in fostering community engagement and enriching online discourse. For example, within online platforms aiming to promote discussion and knowledge-sharing [19, 42], prioritizing the inclusion of data from multiple sources can amplify the richness of conversations and encourage active participation from members. Furthermore, PI systems geared towards facilitating decision-making processes can also benefit substantially from the integration of others' data. For instance, making medical decisions, such as cancer treatment, which often hinges on individuals' perceptions, attitudes, and opinions about specific aspects [9], can be informed by incorporating others' data. The involvement of other's data can help participants reflect on and justify their perceptions of specific situations and treatments, thereby empowering them to make decisions that align with their values and preferences.

5.3 Different mechanism in facilitating reflection

Our findings in Section 4.3 reveal that insights of **confirmation** and **contradiction** within the Recall type were generated exclusively by DP, while insights of **difference** and **similitude** were generated only by NDP. This result suggests a different mechanism of the two data sources in facilitating reflection. DP participants tended to generate insights by comparing their self-cognition with data, thereby recalling their past behaviors and evoking external contexts. In contrast, NDP participants tended to identify others' behaviors captured by data to recall their different or related experiences through comparison. Additionally, our findings in section 4.4 demonstrate insight of **speculation** solely for NDP, indicating an alternative way of making sense of data—interpretation, where individuals use their own experiences to provide plausible explanations for others' data.

Our findings extend the understanding of the anchoring role of one's own data in facilitating reflection [17] and suggest that others' data also serves an anchoring function, albeit through a distinct mechanism—comparison. One's own data enables people to discern conflicts between the "past me" and the "self-recognized me," ultimately fostering a shift in self-understanding. Others' data facilitate reflection by providing a picture of either the "similar me," sharing similar behaviors, or the "different me," displaying different personal lifestyles. Prior literature has emphasized the importance of providing the "right sort of experiences" to foster reflection [68], but many PI tool designs have fallen short in this regard [5]. Our findings indicate that both one's own data challenging self-assumptions and others' data offering relatedness, whether through differences or similarities, serve as effective materials to evoke the "right sort of experiences." However, our data reflection task design only involve simple interaction where people review one single screenshot, which can limit deeper understanding and reflection on personal data. Especially when interpreting others' data, it necessitates an explanation of the underlying contexts [32]. Thus, future research could explore alternative interactions with data, such as speculative methods [46], to enhance individuals' connection to others' data and facilitate the identification of the "similar self."

Moving beyond the conventional one-vs-many comparisons used to define norms [30], our findings also highlight an alternative approach—one-vs-one comparison. This detailed perspective of analyzing data through one-vs-one interactions allows individuals to identify patterns through direct comparison within smaller data sets. Future collaborative PI tools could explore methods to facilitate pair collaboration [32, 36], which is effective in supporting deeper and more spontaneous feedback between individuals.

5.4 Limitation and future work

To size the impact and validity of our study, we identify limitations around three aspects.

First, we recognize that various factors such as occupation, gender, and life conditions can co-influence the relatedness between DP and NDP, thereby potentially affecting reflection and insight generation. Prior research has highlighted that building relatedness is a complex technique that extends beyond merely finding similarities or differences [44]. While our study attempted to facilitate relatedness between participants, we acknowledge the limitations of recruitment due to the feasibility of experiment execution. Future studies can explore identifying key characteristics that better match participants to facilitate building relatedness for collaborative reflection.

Second, the design of the data-reflection task, which involved participants answering reflective questions based on screenshots with word limits (20 and 40 minimum words), may pose limitations in facilitating reflection. Specifically, communication around data—such as inquiry, explanation, and interpretation of underlying contexts—is crucial for fostering reflection but is constrained by presenting only a single screenshot to participants. Additionally, restricting the length of responses could impact the depth of reflection and self-disclosure. Research suggests that longer answers are associated with higher levels of reflection and self-disclosure [12, 31]. Although pilot studies were conducted to balance freedom and constraints in facilitating reflection, there remains a risk that the depth of reflection and self-disclosure could be affected.

Third, we minimised the factors influencing self-disclosure (e.g., personal traits, language, culture [38]) by screening crowd-workers who are English native speakers and grouping them in three age ranges. However, other factors, such as the difference in the enjoyment and the social effect caused by reflecting on one's data and others' data can still influence the degree of self-disclosure [2, 55]. It is challenging to measure and explain reflection and self-exposure separately because of the entangled nature of those factors influencing them. In addition, the frameworks we used to evaluate the level of reflection and self-disclosure are sometimes too subjective and abstract to assess the participants', especially for the disclosure of feelings. We followed a procedure with two coders working independently before aligning to reach a high inter-rater reliability. However, the results remain under the influence of subjective judgments.

6 Conclusion

This paper investigates the different effectiveness of one's own data and others' data in facilitating self-reflection and generating personal insights. Through a crowdsourced approach, we recruited two groups of participants - DP and NDP - to make sense of their own data and others' data through answering opening ended questions in textual format. We evaluate the reflection level, disclosure level, and types of insights derived from the responses of DP and NDP. Our analysis shows that others' data and one's own data have comparable effectiveness in facilitating reflective description (R1), dialogic reflection (R2), and self-disclosure of personal thoughts and feelings. Specifically, we found that one's own data are efficient in supporting the gain of behavioral awareness, while others' data are efficient in helping people express their perceptions, attitudes, and opinions. Furthermore, others' data can serve as an anchor for individuals to recall their own past experiences and trigger the expression of their perceptions, attitudes, and opinions. These results highlight the comparable effectiveness of others' data as protagonistic material in data-informed self-reflection and provide insights into exploiting the advantages while compensating for the shortcomings of the two data sources in collaborative PI for enhancing self-knowledge.

References

[1] Elena Agapie, Lucas Colusso, Sean A Munson, and Gary Hsieh. 2016. Plansourcing: Generating behavior change plans with friends and crowds. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. Association for Computing Machinery, New York, NY, USA, 119–133. https://doi.org/10.1145/2818048.2819943

[2] George Oppong Appiagyei Ampong, Aseda Mensah, Adolph Sedem Yaw Adu, John Agyekum Addae, Osaretin Kayode Omoregie, and Kwame Simpe Ofori. 2018. Examining self-disclosure on social networking sites: A flow theory and privacy perspective. *Behavioral Sciences* 8, 6 (2018), 58.

[3] Bon Adriel Aseniero, Charles Perin, Wesley Willett, Anthony Tang, and Sheelagh Carpendale. 2020. Activity river: Visualizing planned and logged personal activities for reflection. In *Proceedings of the International Conference on Advanced Visual Interfaces*. Association for Computing Machinery, New York, NY, USA, 1–9.

[4] Azy Barak and Orit Gluck-Ofri. 2007. Degree and reciprocity of self-disclosure in online forums. *Cyberpsychology and Behavior* 10, 3 (2007), 407–417. https://doi.org/10.1089/cpb.2006.9938

[5] Eric P.S. Baumer. 2015. Reflective informatics: Conceptual dimensions for designing technologies of reflection. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2015-April. Association for Computing Machinery, New York, NY, USA, 585–594. https://doi.org/10.1145/2702123.2702234

[6] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 5 (2013), 1–27.

[7] Marit Bentvelzen, Julia Dominiak, Jasmin Niess, Frederique Henraat, and Paweł W Woźniak. 2023. How Instructional Data Physicalisation Fosters Reflection in Personal Informatics. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–15.

[8] 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*. Association for Computing Machinery, New York, NY, USA, 1–12.

[9] Andrew BL Berry, Catherine Y Lim, Calvin A Liang, Andrea L Hartzler, Tad Hirsch, Dawn M Ferguson, Zoë A Bermet, and James D Ralston. 2021. Supporting collaborative reflection on personal values and health. *Proceedings of the ACM on human-computer interaction* 5, CSCW2 (2021), 1–39.

[10] Sander Bogers, Janne Van Kollenburg, Eva Deckers, Joep Frens, and Caroline Hummels. 2018. A situated exploration of designing for personal health ecosystems through data-enabled design. In *Proceedings of the 2018 Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 109–120.

[11] Tobias Bornakke and Brian L. Due. 2018. Big–Thick Blending: A method for mixing analytical insights from big and thick data sources. *Big Data and Society* 5, 1 (6 2018). https://doi.org/10.1177/2053951718765026

[12] Matthew Brehmer, Bongshin Lee, Benjamin Bach, Nathalie Henry Riche, and Tamara Munzner. 2017. Timelines Revisited: A Design Space and Considerations for Expressive Storytelling. *IEEE Transactions on Visualization and Computer Graphics* 23, 9 (2017), 2151–2164. https://doi.org/10.1109/TVCG.2016.2614803

[13] Daniel Buschek, Sarah Völkel, Lukas Mecke, Sarah Prange, Clemens Stachl, and Ken Pfeuffer. 2018. Experience sampling as information transmission: Perspective and implications. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. Association for Computing Machinery, New York, NY, USA, 606–611. https://doi.org/10.1145/3267305.3267543

[14] Anna Cherenshchykova and Andrew D Miller. 2019. Family-based sleep technologies: opportunities and challenges. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–6.

[15] Janghee Cho, Tian Xu, Abigail Zimmermann-Niefield, and Stephen Voida. 2022. Reflection in Theory and Reflection in Practice: An Exploration of the Gaps in Reflection Support among Personal Informatics Apps. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 23 pages. https://doi.org/10.1145/3491102.3501991

[16] Eun Kyoung Choe, Bongshin Lee, and M. C. Schraefel. 2015. Characterizing Visualization Insights from Quantified Selfers' Personal Data Presentations. *IEEE Computer Graphics and Applications* 35, 4 (2015), 28–37. https://doi.org/10.1109/MCG.2015.51

[17] 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*. Association for Computing Machinery, New York, NY, USA, 173–182. https://doi.org/10.1145/3154862.3154881

[18] Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1143–1152. https://doi.org/10.1145/2556288.2557372

[19] Chia-Fang Chung, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A Munson. 2017. When personal tracking becomes social: Examining the use of Instagram for healthy eating. In *Proceedings of the 2017 CHI Conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 1674–1687.

[20] Aykut Coşkun and Armağan Karahanoğlu. 2023. Data sensemaking in self-tracking: Towards a new generation of self-tracking tools. *International Journal of Human–Computer Interaction* 39, 12 (2023), 2339–2360.

[21] Andrea Cuttone, Michael Kai Petersen, and Jakob Eg Larsen. 2014. Four data visualization heuristics to facilitate reflection in personal informatics. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8516 LNCS, PART 4 (2014), 541–552. https://doi.org/10.1007/978-3-319-07509-9{_}51

[22] Nediyana Daskalova, Jina Yoon, Yibing Wang, Cintia Araujo, Guillermo Beltran, Nicole Nugent, John McGeary, Joseph Jay Williams, and Jeff Huang. 2020. SleepBandits: Guided Flexible Self-Experiments for Sleep. In *Conference on Human Factors in Computing Systems - Proceedings*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376584

[23] Chris Elsden, David S. Kirk, and Abigail C. Durrant. 2016. A Quantified Past: Toward Design for Remembering With Personal Informatics. *Human-Computer Interaction* 31, 6 (2016), 518–557. https://doi.org/10.1080/07370024.2015.1093422

[24] Chris Elsden, Bettina Nissen, Andrew Garbett, David Chatting, David Kirk, and John Vines. 2016. Metadating: exploring the romance and future of personal data.

In *Proceedings of the 2016 chi conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 685–698.

[25] Daniel A Epstein, Alan Borning, and James Fogarty. 2013. Fine-grained sharing of sensed physical activity: A value sensitive approach. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. Association for Computing Machinery, New York, NY, USA, 489–498.

[26] Daniel A Epstein, Bradley H Jacobson, Elizabeth Bales, David W McDonald, and Sean A Munson. 2015. From" nobody cares" to" way to go!" A Design Framework for Social Sharing in Personal Informatics. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. Association for Computing Machinery, New York, NY, USA, 1622–1636.

[27] Daniel A Epstein, Siyun Ji, Danny Beltran, Griffin D'Haenens, Zhaomin Li, and Tan Zhou. 2020. Exploring design principles for sharing of personal informatics data on ephemeral social media. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–24.

[28] 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*. Association for Computing Machinery, New York, NY, USA, 731–742.

[29] Irwin Feinberg. 1974. Changes in sleep cycle patterns with age. *Journal of psychiatric research* 10, 3-4 (1974), 283–306.

[30] 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 (2018), 1–21.

[31] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on reflection: Framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*. Association for Computing Machinery, New York, NY, USA, 216–223. https://doi.org/10.1145/1952222.1952269

[32] Mikhaila Friske, Jordan Wirfs-Brock, and Laura Devendorf. 2020. Entangling the roles of maker and interpreter in interpersonal data narratives: Explorations in yarn and sound. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 297–310.

[33] Jeana Frost, Michael Massagli, et al. 2008. Social uses of personal health information within PatientsLikeMe, an online patient community: what can happen when patients have access to one another's data. *Journal of medical Internet research* 10, 3 (2008), e1053.

[34] Katerina Gorkovenko, Dan Burnett, Murray-Rust, James Thorp, and Daniel RIichards. 2019. Supporting Real-Time Contextual Inquiry through Sensor Data. *Ethnographic Praxis in Industry Conference Proceedings* 2019, 1 (2019), 554–581. https://doi.org/10.1111/1559-8918.2019.01307

[35] Rebecca Gulotta, Alex Sciuto, Aisling Kelliher, and Jodi Forlizzi. 2015. Curatorial agents: How systems shape our understanding of personal and familial digital information. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3453–3462.

[36] Jo E Hannay, Tore Dybå, Erik Arisholm, and Dag IK Sjøberg. 2009. The effectiveness of pair programming: A meta-analysis. *Information and software technology* 51, 7 (2009), 1110–1122.

[37] Jina Huh, Leslie S Liu, Tina Neogi, Kori Inkpen, and Wanda Pratt. 2014. Health vlogs as social support for chronic illness management. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–31.

[38] Emmi Ignatius and Marja Kokkonen. 2007. Factors contributing to verbal self-disclosure. *Nordic Psychology* 59, 4 (1 2007), 362–391. https://doi.org/10.1027/1901-2276.59.4.362

[39] Simon L Jones and Ryan Kelly. 2018. Dealing with information overload in multifaceted personal informatics systems. *Human–Computer Interaction* 33, 1 (2018), 1–48.

[40] Matthew Jörke, Yasaman S Sefidgar, Talie Massachi, Jina Suh, and Gonzalo Ramos. 2023. Pearl: A Technology Probe for Machine-Assisted Reflection on Personal Data. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. Association for Computing Machinery, New York, NY, USA, 902–918.

[41] Gyuwon Jung, Sangjun Park, and Uichin Lee. 2024. DeepStress: Supporting Stressful Context Sensemaking in Personal Informatics Systems Using a Quasi-experimental Approach. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–18.

[42] Kasper Karlgren, Barry Brown, and Donald McMillan. 2022. From Self-Tracking to Sleep-Hacking: Online Collaboration on Changing Sleep. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–26.

[43] Kasper Karlgren and Donald McMillan. 2022. Designing for Extreme Sleepers: Rethinking the Rhythms of Sleep Technology. In *Nordic Human-Computer Interaction Conference*. Association for Computing Machinery, New York, NY, USA, 1–17.

[44] Maria Karyda, Elisa D. Mekler, and Andres Lucero. 2021. Data agents: Promoting reflection through meaningful representations of personal data in everyday life. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 11 pages. https://doi.org/10.1145/3411764.3445112

[45] Laura Koesten, Kathleen Gregory, Paul Groth, and Elena Simperl. 2021. Talking datasets–understanding data sensemaking behaviours. *International journal of human-computer studies* 146 (2021), 102562.

[46] Albrecht Kurze, Andreas Bischof, Sören Totzauer, Michael Storz, Maximilian Eibl, Margot Brereton, and Arne Berger. 2020. Guess the Data: Data Work to Understand How People Make Sense of and Use Simple Sensor Data from Homes. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376273

[47] Lan Li, Dey, and Jodi Forlizzi. 2010. *A Stage-Based Model of Personal Informatics Systems*. Association for Computing Machinery, New York, NY, USA. 2642 pages.

[48] Gianina Luca, José Haba Rubio, Daniela Andries, Nadia Tobback, Peter Vollenweider, Gérard Waeber, Pedro Marques Vidal, Martin Preisig, Raphaël Heinzer, and Mehdi Tafti. 2015. Age and gender variations of sleep in subjects without sleep disorders. *Annals of medicine* 47, 6 (2015), 482–491.

[49] Deborah Lupton. 2020. *Data selves: More-than-human perspectives*. Polity Cambridge, Cambridge, UK.

[50] Alina Lushnikova, Kerstin Bongard-Blanchy, and Carine Lallemand. 2022. What Aspects of Collaboration are Meaningful to You? Informing the Design of Self-Tracking Technologies for Collaboration. In *Adjunct Proceedings of the 2022 Nordic Human-Computer Interaction Conference*. Association for Computing Machinery, New York, NY, USA, 1–5.

[51] Lena Mamykina, Andrew D Miller, Catherine Grevet, Yevgeniy Medynskiy, Michael A Terry, Elizabeth D Mynatt, and Patricia R Davidson. 2011. Examining the impact of collaborative tagging on sensemaking in nutrition management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 657–666.

[52] JC Marquie and J Foret. 1999. Sleep, age, and shiftwork experience. *Journal of sleep research* 8, 4 (1999), 297–304.

[53] Nick Merrill, John Chuang, and Coye Cheshire. 2019. Sensing is believing: What people think biosensors can reveal about thoughts and feelings. In *Proceedings of the 2019 on Designing Interactive Systems Conference*. Association for Computing Machinery, New York, NY, USA, 413–420.

[54] Jimmy Moore, Pascal Goffin, Jason Wiese, and Miriah Meyer. 2021. Exploring the Personal Informatics Analysis Gap:"There's a Lot of Bacon". *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 96–106.

[55] Samar Mouakket. 2019. Information self-disclosure on mobile instant messaging applications: Uses and gratifications perspective. *Journal of Enterprise Information Management* 32, 1 (2019), 98–117. https://doi.org/10.1108/JEIM-05-2018-0087

[56] San San Nguyen, Da-jung Kim, Ting Miao, and Yaliang Chuang. 2020. Designing for Triggering Self-Investigations and Reflections on Factors Related to Sleep Health. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. Association for Computing Machinery, New York, NY, USA, 1–4.

[57] Giovanna Nunes Vilaza, Raju Maharjan, David Coyle, and Jakob Bardram. 2020. Futures for Health Research Data Platforms From the Participants' Perspectives. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*. Association for Computing Machinery, New York, NY, USA, 1–14.

[58] Laura Pina, Sang-Wha Sien, Clarissa Song, Teresa M Ward, James Fogarty, Sean A Munson, and Julie A Kientz. 2020. DreamCatcher: exploring how parents and school-age children can track and review sleep information together. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW1 (2020), 1–25.

[59] Aare Puussaar, Adrian K Clear, and Peter Wright. 2017. Enhancing personal informatics through social sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 6936–6942.

[60] Aare Puussaar, Adrian K. Clear, and Peter Wright. 2017. Enhancing personal informatics through social sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 6936–6942. https://doi.org/10.1145/3025453.3025804

[61] Amon Rapp and Maurizio Tirassa. 2017. Know Thyself: A Theory of the Self for Personal Informatics. , 335–380 pages. https://doi.org/10.1080/07370024.2017.1285704

[62] Ruth Ravichandran, Sang-Wha Sien, Shwetak N Patel, Julie A Kientz, and Laura R Pina. 2017. Making sense of sleep sensors: How sleep sensing technologies support and undermine sleep health. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 6864–6875.

[63] Susan Redline, H Lester Kirchner, Stuart F Quan, Daniel J Gottlieb, Vishesh Kapur, and Anne Newman. 2004. The effects of age, sex, ethnicity, and sleep-disordered breathing on sleep architecture. *Archives of internal medicine* 164, 4 (2004), 406–418.

[64] Saeyoung Rho, Injung Lee, Hankyung Kim, Jonghyuk Jung, Hyungi Kim, Bong Gwan Jun, and Youn-kyung Lim. 2017. Futureself: what happens when

we forecast self-trackers? Future health statuses?. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. Association for Computing Machinery, New York, NY, USA, 637–648.

[65] Minna Ruckenstein. 2014. Visualized and Interacted Life: Personal Analytics and Engagements with Data Doubles. *Societies* 4, 1 (2014), 68–84. https://doi.org/10.3390/soc4010068

[66] Herman Saksono, Carmen Castaneda-Sceppa, Jessica Hoffman, Magy Seif El-Nasr, Vivien Morris, and Andrea G Parker. 2019. Social reflections on fitness tracking data: A study with families in low-SES neighborhoods. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14.

[67] Roger C Schank and Robert P Abelson. 2013. *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Psychology Press, Hillsdale, New Jersey.

[68] Petr Slovak, Chris Frauenberger, and Geraldine Fitzpatrick. 2017. Reflective practicum: A framework of sensitising concepts to design for transformative reflection. *Conference on Human Factors in Computing Systems - Proceedings* 2017-May (2017), 2696–2707. https://doi.org/10.1145/3025453.3025516

[69] Peter Tolmie, Andy Crabtree, Tom Rodden, James Colley, and Ewa Luger. 2016. This has to be the cats - Personal data legibility in networked sensing systems. *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW* 27 (2016), 491–502. https://doi.org/10.1145/2818048.2819992