Thinking Outside the Data Box: Investigating the Potential of Data Manipulation for Self-Reflection on Personal Data

Yeohyun Jung
KAIST
Republic of Korea
yeohyun@kaist.ac.kr

Kwangyoung Lee
KAIST
Republic of Korea
kwangyoung@kaist.ac.kr

Hwajung Hong
KAIST
Republic of Korea
hwajung@kaist.ac.kr

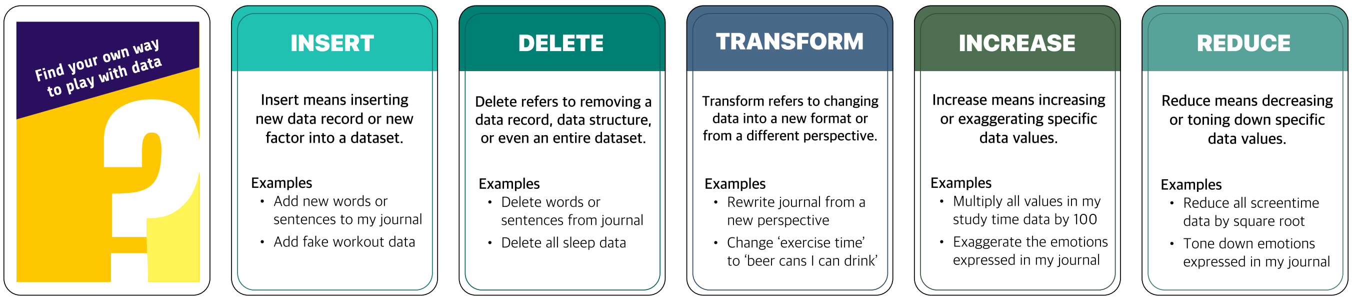


Figure 1: Data manipulation is an approach that involves altering the value or structure of self-tracking data. With five data manipulation typology (Insert, Delete, Transform, Increase, and Reduce), we applied data manipulation through exploratory workshops and a one-week field trial. The above cards, which served as key materials for this study, provide an overview of the different types of data manipulation, along with definitions and examples for each type.

ABSTRACT

In the practice of personal informatics (PI), self-reflection is crucial for enhancing self-knowledge and driving behavior change. Numerous studies have focused on effectively interpreting and representing data to support self-reflection. However, despite their efforts, some self-trackers find themselves stuck in repetitive insights and stagnant process. For them, a fundamental shift beyond re-representing existing data could provide a significant opportunity. We explore data manipulation—altering the data value or structure—as an alternative approach. We conducted an exploratory workshop and a one-week field trial with 10 self-trackers, using five types of data manipulation. We found that data manipulation could revitalize self-reflection, uncovering diverse perspectives and overlooked aspects. It also fostered positive illusions and emotions, potentially setting the stage for behavioral change and engagement. However, it introduces perceptual distortions and has limited applicability, highlighting the importance of balanced use. We further discuss design implications for integrating data manipulation into future PI systems.

CCS CONCEPTS

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

KEYWORDS

Data manipulation; Self-tracking Data; Self-reflection; Stagnation; Personal informatics

ACM Reference Format:
Yeohyun Jung, Kwangyoung Lee, and Hwajung Hong. 2025. Thinking Outside the Data Box: Investigating the Potential of Data Manipulation for Self-Reflection on Personal Data. In *Designing Interactive Systems Conference (DIS '25), July 5–9, 2025, Funchal, Portugal.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3715336.3735840

1 INTRODUCTION

Taylor had been diligently self-tracking his workouts with the smartwatch for over a year. However, with the same routines, goals, and time slots repeating day after day, Taylor found it increasingly difficult to extract anything meaningful from the data, during reflection. The fitness progress also seemed to have plateaued. "I've been working so hard to collect and reflect on this data, but if I'm not gaining anything new, does self-tracking even mean anything to me anymore? Should I just stop?" Taylor wondered. Taylor tried various approaches: visualizations to present his data in various ways and asking an AI for personalized insights… But these efforts only yielded similar reflections and failed to bring any real change. Then, one sleepless night, a bold thought crossed the mind: What if I manipulate the data itself? Taylor decided to manipulate the workout data. Breaking the box of original data felt strange yet thrilling, unlocking a new perspective.

** This scenario is based on the real concerns and experiences shared by one of the participants in this study, referred to here by the pseudonym Taylor.*

DIS '25, July 5–9, 2025, Funchal, Portugal

ACM ISBN 979-8-4007-1485-6/2025/07
https://doi.org/10.1145/3715336.3735840

Self-tracking refers to the process where individuals collect and reflect on data about themselves to gain self-knowledge and facilitate positive behavior change [41]. It involves the practice of self-reflection that allows users to generate insights, improve behaviors, and refine their self-concept [22, 42]. To effectively reflect on their data, self-trackers leveraged visualizations, both quantitative and qualitative analyses [14]. They discover meaningful insights, identify goals, and make informed decisions how to achieve the goal from the data [34, 41, 53].

However, having meaningful and supportive reflection throughout the self-tracking journey is often a significant challenge. A critical issue identified in Human-computer interaction (HCI) and PI research is that many self-trackers encounter a state of *"stagnation"* in their reflective practices [12, 17, 23]. This state arises despite their efforts to engage thoughtfully with their data, leaving them unable to derive new insights or experience meaningful progress anymore. This often stems from the decreased interest and novelty in the data, feeling as though there are no further insights to uncover [24, 40, 72]. Self-trackers also feel stuck when they cannot make meaningful behavioral changes [24]. The emotional toll of negative thoughts and emotions aroused from the data further hinder constructive reflection and growth. Rather, it turns the spiral into a cycle of negative thoughts and emotions [16, 65]. Eventually, this stagnation in the self-reflection process can demotivate self-tracking and lead to lapse (i.e., pausing or abandoning self-tracking altogether) [25].

Over the years, researchers have explored various approaches to support reflection in order to enhance self-knowledge and sustain motivation for behavior change. One prominent method is visualization, which seeks to present existing data in ways that are more intuitive and insightful [13, 20, 43, 52, 62]. By employing techniques such as the framing effect [68], visualization researchers have designed tools that help users see their data in new and fresh representations. AI-driven approaches have emerged as a means of supporting reflection. Interpretations from AI [64] or AI-based conversational agent [38] could provide accurate and personalized reflections. Social components have also proven effective, where individuals compare their self-tracking data within social relationships and exchange feedback, fostering collaborative reflection [26, 46, 61].

While these approaches excel at presenting and conveying data in more compelling ways, they primarily focus on reframing the same data to support reflection. However, individuals stuck in their self-tracking journey face a different challenge: they are saturated and struggle to make further progress. Breaking through this stagnation requires fresh input that moves beyond merely re-representing existing data. We draw inspiration from "story editing", a technique commonly used in psychotherapeutic settings. It involves intentionally altering existing personal narratives to foster constructive thoughts and new perspectives, ultimately leading to significant emotional and behavioral changes [9]. In line with prior studies investigating ways of engaging in data, Fleck and Fitzpatrick [27] also presented a conceptual framework through which reconfiguring data can deconstruct existing understandings and knowledge. Therefore, we considered an approach that goes beyond re-representing data to directly engage with its properties, values, and structure—referring to "Data Manipulation". This brings us to the key research question: *Could manipulating the data itself unlock new opportunities for exploration and self-reflection?*

To address this question, we explore data manipulation's potential positive and negative impacts on self-tracking. After organizing the data manipulation typology (i.e., Insert, Delete, Transform, Increase, Reduce), we conducted exploratory workshop where participants generated and developed ideas for data manipulation, followed by a one-week field trial in which they repeatedly applied data manipulation to their self-tracking. With 10 self-trackers who felt saturated with self-tracking, our study aims to investigate how self-trackers perceive the approach of manipulating data and the experiences and impacts it produces. Our findings indicated that the participants found data manipulation intriguing and came up with various approaches to apply data manipulation. Data manipulation could reinvigorate and enrich self-reflection by helping individuals gain deeper insights and see themselves from more diverse perspectives. This process also fosters a constructive mindset for an action, opening possibilities for sustained behavioral engagement and change. However, it also introduced challenges such as distorted perceptions and the reality that its effects are not uniform across all self-trackers or data domains. This highlighted the need for balanced and nuanced usage of data manipulation to leverage its potential. We propose several design implications for integrating data manipulation into future PI systems based on these findings.

The key contributions of this study are threefold:

- We applied data manipulation as an alternative approach, which is adjusting the data itself, to explore its potential and role for self-tracking practices with a focus of self-reflection.
- We report empirical findings on perception, applications, experiences, and impacts of data manipulation.
- We discuss design implications for future PI systems that incorporate data manipulation.

2 RELATED WORKS

This section provides an overview of previous research about supporting self-reflection and challenges in self-reflection, with a focus of stagnation. We then review the approaches of altering data.

2.1 Challenges and Stagnation in Self-Reflection

In the field of personal informatics, self-reflection is a critical practice that supports behavior change and self-improvement [42]. Additionally, as Fleck and Fitzpatrick [27] outlined in their framework of five reflection levels—R0: descriptive reflection, R1: explanatory reflection, R2: dialogic reflection, R3: transformative reflection, and R4: critical reflection—self-reflection is recognized as a multifaceted practice involving different depths and dimensions. However, achieving productive and insightful reflection often remains as one of the most significant challenges [14]. Some individuals encounter obstacles in self-reflection because they feel difficulties from the beginning. They struggle to comprehend the data they have collected and find it difficult to derive actionable insights [37]. According to Alqahtani et al. [2], this state of "uncertainty" arises when individuals are unable to fully understand and make progress upon the collected data.

In addition to those who do not know how to reflect on their data, there are some individuals experiencing stagnation after engaging in self-reflection. They believe they have already learned everything they can about themselves, and they feel that there are no new and exciting insights to uncover [23, 24, 72]. Moreover, the standstill in creating meaningful behavior change from their data leaves them feeling frustrated, which weakens their desire to continue with self-tracking [24]. Lastly, reflection of negative self-tracking data can provoke intense negative emotions, which can prevent people from proceeding with self-reflection, even if they are aware of how to do it [16, 65]. For example, individuals with eating disorders [17] or multiple chronic conditions [3] often experience stress and hopelessness when faced with their data, which makes it difficult to move forward with productive and constructive reflections.

This study focuses on individuals who are stuck in the self-reflection process. To help them overcome this stagnation and foster meaningful and engaging self-reflection, we aim to explore a shift that encourages them to interact with their data more actively and creatively, and examine the impact of this approach on their self-reflection.

2.2 Supporting Self-reflection in Personal Informatics

Many studies have explored various approaches to support better self-reflection. Among these, personalized visualizations have emerged as a field aimed at representing data into personally relevant form in order to encourage deriving actionable insights and behavior change [13, 30, 43, 52]. Strategies involve visualizing personal data through dashboards, enabling simple interactions to explore data [32, 62], and framing or filtering data for better comprehension [68]. For instance, Epstein et al. [20] introduced "visual cut," which refers to filtering specific subsets of collected data, and explored how different visualizations can support users in making sense of their tracked data. Additionally, there has been growing interest in AI-driven data interpretation. AI systems could analyze collected data to generate insights and offer users constructive interpretations [38, 64]. Collaborative reflection within social relationships is another well-explored approach, where individuals reflect on their data with family members [61], colleagues [46], or others with data similar to their own [26]. These studies have shown that sharing and comparing data in a social and collaborative setting can enable users to reflect more deeply and uncover new insights about their behaviors and experiences. With advancements in large language models (LLMs), users can now collaborate with LLMs during the self-reflection process, discussing their data with AI systems to achieve deeper reflection [64].

However, these previous studies has primarily focused on helping users interpret and reflect on the existing data more effectively. For self-trackers who are stagnated and saturated with self-reflection, interpreting and re-representing their existing data may yield limited insights and impacts. Several theories on learning and behavioral change suggest that foundational and innovative approaches, which differ significantly from prior actions, are essential for driving meaningful transformation [4, 48]. Moreover, Rapp and Tirassa [58] emphasized that while current personal informatics systems often aim to provide optimal interpretations of user data, it could be a valuable opportunity for self-trackers to explore and interact with their data in diverse ways, referring to *user-controlled proliferation of takes on the user's data*. Therefore, this study aims to investigate an alternative approach, enabling self-trackers to explore and engage with their data in various ways.

2.3 Approaches to Altering Data

While much of the prior research has focused on interpreting, visualizing, or representing existing data, some studies have explored approaches that involve modifying the data itself. Changing data is commonly utilized in psychological therapy since it can help with the cognitive reframing of previous thoughts and beliefs [59]. One such technique named "Story Editing" involves transforming an individual's personal narrative into a different story, helping them reshape their perception of themselves and their past experiences [9]. This technique disrupts destructive thought patterns and encourages a more positive perspective [50]. In both HCI and PI fields, there have been studies exploring the modification and manipulation of data. Epstein et al. [21] introduced an interactive modification approach that allows users to adjust their detailed step data before sharing data, as a value-sensitive approach. By removing private data and replacing it with more realistic values, this approach alleviated concerns about privacy intrusion. Moreover, when discrepancies arose between a user's perception and the collected data, reducing the gap by actively transforming and restructuring the data helped improve self-reflection [7, 19].

While previous studies have primarily focused on using data manipulation as a tool for privacy protection or resolving discrepancies between data and personal perceptions, the potential for data manipulation in self-tracking could go beyond this. Baumer and Fitzpatrick [27] introduced the concept of "Reflective Informatics," which emphasizes that users can gain new insights by questioning and reconstructing their existing data. They highlight that manipulating data can disrupt prior perceptions, enabling users to reinterpret their experiences and drive further transformation. Building upon this concept, our study aims to explore how active and creative data manipulation can facilitate the discovery of new insights and impact self-reflection and behavior change in self-tracking. Therefore, by applying data manipulation to individuals who are stuck in the self-reflection process, we sought to investigate how they generate new insights, the effects on their self-reflection and subsequent behaviors, and the overall experience.

3 METHOD

Under the concept of "Data manipulation," which refers to changing data value or structure, this study aims to explore the approaches, opportunities, and challenges associated with the concept. To achieve this, we conducted exploratory workshops to gather participants' thoughts and ideas through creative engagement. Following this, we carried out a one-week field trial where individuals repeatedly applied data manipulation for a week in order to investigate experiences and the impacts of data manipulation on the self-reflection.

3.1 Participants

We introduced the approach of data manipulation with the observation that self-tracking data is typically not altered after collection. Based on this, the study was advertised as an engaging activity

ID	Gender	Age	Data Domain	Duration
P1	M	25	Physical Activity	2 years
			Sleep	
			Journaling	
P2	F	28	Physical Activity	2.6 years
			Finance	
			Productivity (Lapse and Resume)	
P3	F	24	Physical Activity	3 years
			Smartphone Screentime	3 years
			Finance	1 year
P4	M	20	Physical Activity	3 years
			Finance	2 years
			Work Hours	7 months
P5	M	24	Finance	5 years
P6	F	29	Physical Activity	5 years
			Productivity (Lapse and Resume)	
P7	M	33	Physical Activity	2 years
			Productivity	4 years
P8	M	28	Finance	5 years
P9	F	23	Physical Activity (Lapse and Resume)	5 years
			Journaling	
P10	M	23	Journaling	2 months

Table 1: The table presents participants' gender, age, self-tracking data domains, and duration of each self-tracking activity. Self-tracking data that participants reported as having been lapsed and resumed is labeled as (Lapse and Resume).

where people can playfully and creatively interact with their self-tracking data. Any individuals who 1) had prior experience with self-tracking, and 2) had interest in manipulating their self-tracking records were invited. The study event was advertised through social media and online school community posts. We enrolled 10 participants (M = 25.70, SD = 3.77, 4 females) who were able to attend both the exploratory workshop and one-week field trial. Through a pre-survey, we gathered demographic information as well as details about each participant's self-tracking habits, including the areas they track, the duration of their tracking, and the tools they use. They received 60,000 KRW (approximately $45 USD) for their participation. The study was approved by the Institutional Review Board (IRB) of the researcher's institution. Table 1 presents the demographic and self-tracking information of the participants.

3.2 Exploratory Workshop

The approach to manipulating data itself could have been relatively unfamiliar for the participants. It was necessary to create a setting that would serve as a generative phase, enabling participants to imagine and articulate their thoughts on the concept. Workshop is known as an effective methodology to support people to meaningfully engage with the speculative concept and elicit their ideas through creative interaction [55]. Through the workshops, we aimed to observe how they perceive data manipulation and brainstorm its potential applications.

3.2.1 Data Manipulation Typology. To facilitate more creative data manipulation by participants throughout the study, we defined five types of data manipulation. These types were inspired by the Data Manipulation Language (DML) which is a set of commands used to manipulate data in a database when performing data preprocessing [29]. We adapted three key techniques from DML, which are Insert, Delete, and Update [66]. Insert refers to adding new data records or factors to the dataset, while delete involves removing existing data records or factors [70]. Update refers to modifying existing data records with new ones [70]. We thought that update implies a variety of actions such as changing the data value or its form. To minimize the confusion that participants can feel from the broad concept, we divided it into three specific actions: Increase, Reduce, and Transform. Increase refers to increasing or exaggerating the data value. Reduce means decreasing or diminishing the data value, and when the structure of the data is changed beyond simple value modifications, it is categorized as transform. Our typology also drew insights from prior work in the field of PI. Epstein et al. [21] proposed an interactive data transformation system for privacy preservation, which included three core operations: Add, Remove, and Shift. Add involved increasing or inserting data values, Remove referred to eliminating data by reducing values to zero, and Shift reconfigured the temporal structure of the data. Moreover, users had the flexibility to modify their data, even if it differed from reality, rather than correcting inaccurate data since the goal of data transformation was privacy preservation. Inspired by this, we also

Types	Description	Example
Insert	**Original:** Adding missing data records and new factors to the dataset	Adding exercise data that were missed due to forgetting to wear a wearable device
	Extended: Intentionally adding false data records and factors to the dataset	Falsely adding an exercise data for a workout that did not occur
Delete	**Original:** Removing incorrectly recorded data, duplicate data, noise, or outliers from the dataset	Deleting an abnormally high expense that skews the average
	Extended: Deleting data that one wishes to hide based on the subjective thoughts and feelings	Deleting a paragraph from a journal that recalls negative emotions
Transform	**Original:** Converting data into a format suitable for analysis	Standardizing the units in a dataset
	Extended: Intentionally distorting or altering data in a false manner to change its original meaning	Rewriting a negatively worded journal in a positive way
Increase	**Original:** Increasing the value of incorrectly recorded data	Increasing the recorded exercise time to reflect the reality
	Extended: Exaggerating the value of data beyond what is true	Recording 2 hours of study time when only 1 hour was actually spent
Reduce	**Original:** Reducing unnecessary parts of data and decreasing the value of incorrectly recorded data	Reducing the recorded phone usage time to reflect the reality
	Extended: Intentionally decreasing the value of data	Falsely report a 1000 kcal intake as 800 kcal

Table 2: This table contains descriptions and examples of each type of data manipulation. Five types of data manipulation were established, and each definition was further expanded to create more creative and flexible manipulations.

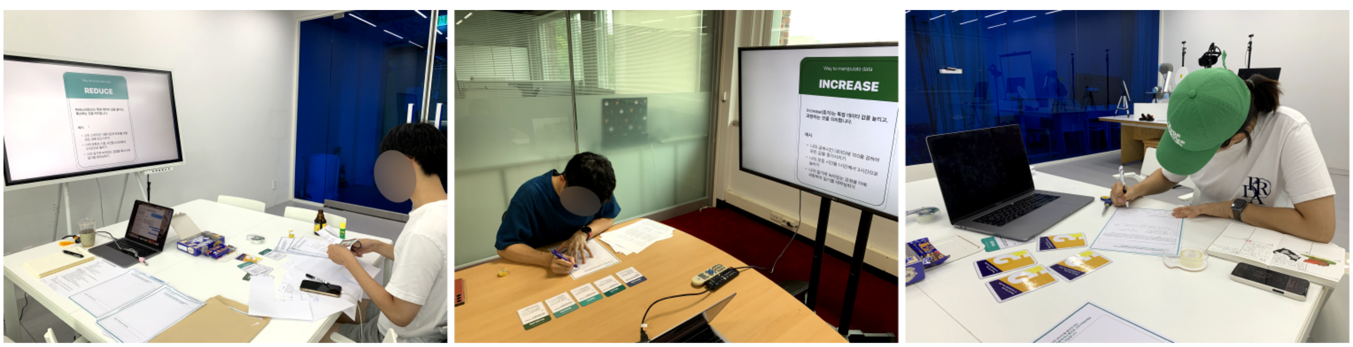


Figure 2: Capturing the moments during the exploratory workshop.

expanded the interpretation of five types—Insert, Delete, Transform, Increase, and Reduce—beyond their original meaning in the context of data preprocessing to form a broader and more creative interaction. This extended interpretation would inspire users to explore diverse ways to manipulate their data. Each description and illustrative examples are organized in the table 2.

3.2.2 Material. Based on the data manipulation typology, we designed cards that included the definition of each type and examples of its application (see Figure 1). Modeled after a card game, the workshop was designed to foster a comfortable and lively atmosphere. Additionally, the random flipping of these cards was intended to minimize any potential order or interaction effects between the types.

For each type, we also created five A4-sized paper worksheets to help participants express their ideas and redefine their self-tracking records (see Figure 3). We posited that paper worksheets could foster flexible and creative data manipulation activities, as they allow participants to express and externalize a wide range of ideas. Furthermore, engaging with data in a hands-on, manual manner through paper would encourage more active and enjoyable participation, drawing on insights from prior research on manual engagement in self-tracking [1, 6]. Each worksheet was divided into two sections: the top section allowed participants to record or print their original self-tracking data and reshape it according to the selected type, while the bottom section provided space for participants to explain the reasoning behind their data manipulation.

3.2.3 Procedure. Among the 10 participants, two groups of three members, who already knew each other, applied to participate as a team. For the remaining four participants, we attempted to form a group, but scheduling conflicts made it difficult for them to meet at the same time. Instead, the researchers conducted the workshop with these four participants individually. During these sessions, researchers closely observed and facilitated their activities to encourage the creative engagement and the generation of divergent ideas, by actively asking follow-up questions.

Prior to the workshop, participants were guided to bring their own self-tracking data, either by bringing the self-tracking tool where their data was stored, or by printing a copy of their self-tracking data. Most participants brought their phones, laptops, or smartwatches to view their self-tracking data, while one participant extracted the data from the self-tracking app in Excel format and brought printed copies. After attending the workshop, participants reviewed and signed the consent form, confirming their understanding, and were informed they could withdraw at any time without consequences. The session was audio-recorded with their permission.

We started the first session with ice-breaking and a brief interview about their self-tracking practices, given that every participant had prior self-tracking experiences. We asked participants about their experiences and challenges in reflecting on their data, as well as any further interactions with data they had attempted to overcome the reflection difficulties. For those who had experienced lapses, we further discussed their reasons for lapse and motivations for restarting.

In the second session, we introduced the typology of manipulating data using the workshop cards. We explained the definition and examples of each type, making sure participants understood the concept of data manipulation. Then, we distributed the workshop cards and paper worksheets for the activity. Participants were asked to randomly select one of the five cards by flipping them over. They then closely referred to their self-tracking data—either through their personal self-tracking tools or printed copies—to find portions of the data they found interesting to manipulate. Next, they manually transferred the raw data or a structured version of it (e.g., tables, lists) onto the paper worksheet. For participants who brought textual data such as journal entries, we provided a printer because handwriting long passages would have been burdensome. These participants attached the printed excerpts to the paper worksheet and proceeded with the manipulation activity. They came up with various application ideas, freely organizing and articulating their thoughts on the worksheet (see Figure 3). Researchers kept observing how the participants applied each type of data manipulation and actively prompted their engagement by asking follow-up questions about their activities.

At the end of the session, we asked participants about their overall workshop experience. The questions explored how they perceived the concept of manipulating data during the workshop, any changes in their perspectives on interacting with PI data, and any anticipated impact of applying the concept of data manipulation on self-reflection. The workshop lasted between 66 and 97 minutes.

3.3 One-week Field Trial

While the exploratory workshops focused on generating and discussing data manipulation ideas, the one-week field trial offered a tangible opportunity to put these ideas into practice. One-week field trial was also designed to help participants refine and adapt their ideas to their individual self-tracking activities. In this study, we aimed to capture the resulting impacts and experiences of they gained through the week-long process.

3.3.1 Material. Each participant received seven A4-sized paper worksheets to document their data manipulation activities during a week (see Figure 4). The worksheets were divided into two sections: the top allowed participants to record or attach their original self-tracking data on the left and reshape it on the right. This design enabled participants to view the original and manipulated data side by side, helping them reflect on the difference between original and manipulated data. The bottom section included four questions aimed at understanding participants' reasons for manipulating their data, their experience of making these changes, how the changes made the previous day influenced the current day, and how they were expected to impact the next day's self-tracking. The last two questions focused on assessing the ongoing effects of these changes on participants' achievement.

3.3.2 Procedure. We found that preferences and applicability regarding types of data manipulation varied depending on participants' needs and the domain of their self-tracked data. Taking this into account, participants were guided to select the one type they found most relevant to their self-tracking practices, rather than being required to apply all types over the week. However, they were encouraged to contact the researcher if they wished to try an additional type for more flexible data interactions or if they encountered challenges with their chosen type. After selecting their preferred type, participants attended an introductory session. They were instructed to complete a daily worksheet to track their data, reshape it, and answer four reflective questions. At the end of the week, participants participated in a 30-minute post-interview, conducted either in person or online, to discuss their experiences, opportunities, and challenges with manipulating their data. All interviews were audio-recorded with informed consent.

3.4 Data Analysis

This study generated valuable data including 120 paper worksheets and about 780 minutes of audio recordings from exploratory workshops and interviews. The audio recordings were transcribed after all identifiable information was removed to ensure anonymity.

The primary pillars of analysis were participants' perceptions and reactions to data manipulation, the ways they applied it, the experiences and impacts of data manipulation with a focus of self-reflection and subsequent actions. Two lead researchers analyzed both the interview transcripts and worksheets using thematic analysis [10]. Starting from the initial codebook, researchers conducted the iterative process of open coding to align with the objective and primary focus of this study. Our initial codebook included the following categories: (1) Interpretation and Application, capturing how the participants understand the concept of data manipulation and generate ideas for its application; (2) Motivations for Data Manipulation, focusing on the reasons the participants apply data manipulation; (3) Self-reflection, highlighting the impact of data manipulation on reflective practices; (4) Behavioral Transformation, addressing the effect of data manipulation on the behaviors; (5) Emotions and Experiences, exploring the emotions and user

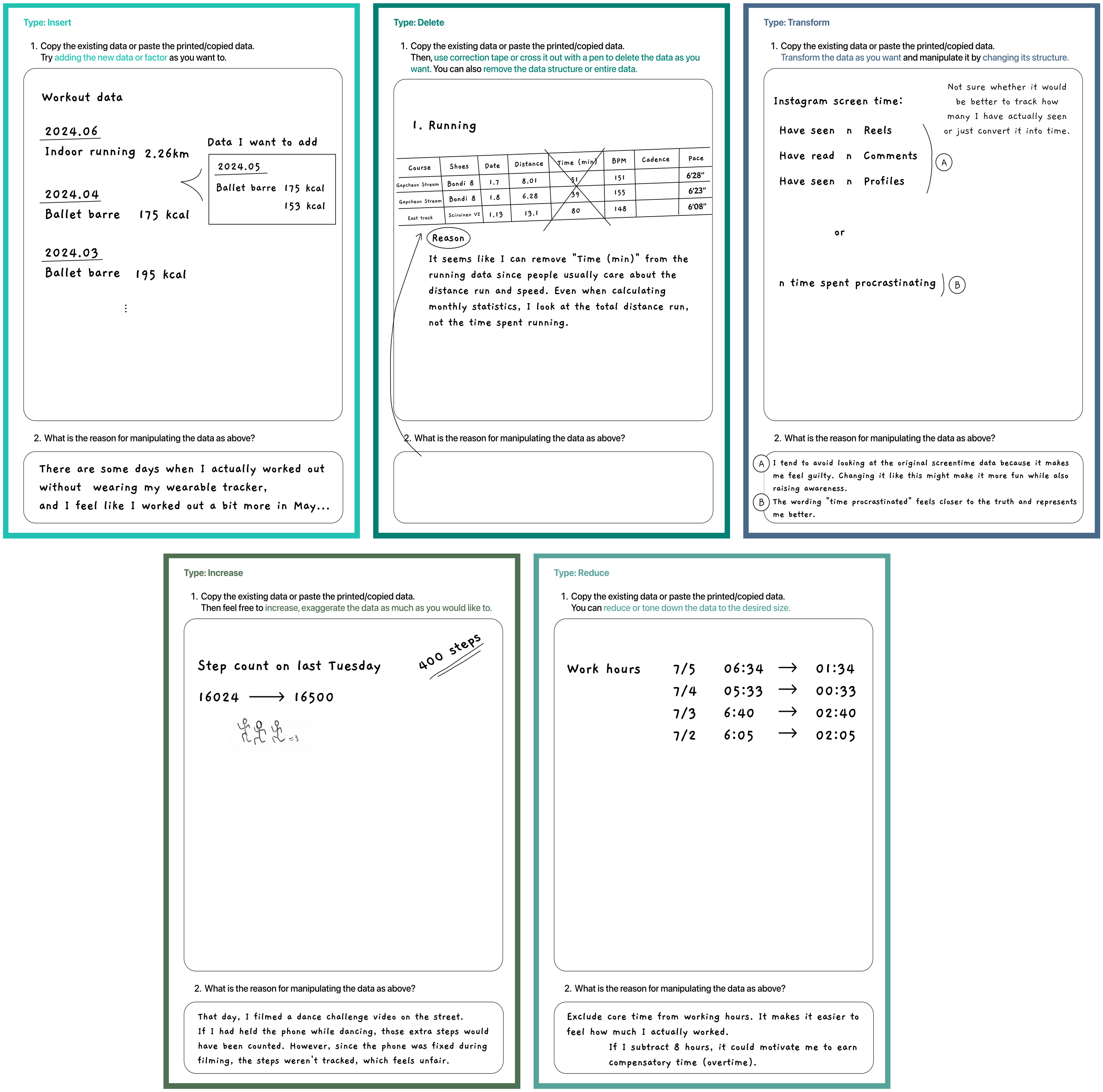
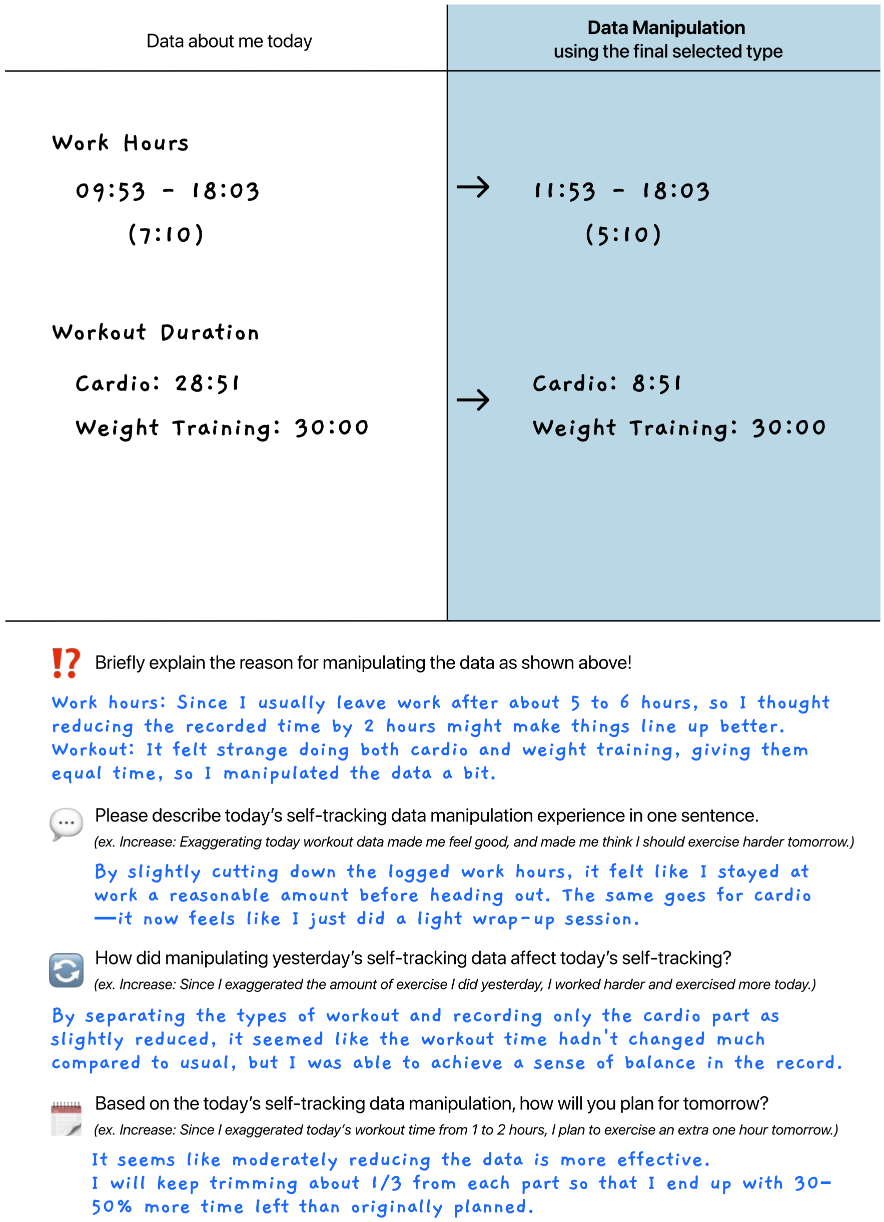


Figure 3: Translated Worksheet examples generated by the participants. Participants, who selected a data manipulation type from the cards, transferred their self-tracking data onto the corresponding paper worksheet. They then manipulated the data's values or structure. They explained the reasons and motivations for making these manipulation. The original worksheets are attached as Figure 5 in the Appendix A.

experiences associated with data manipulation; and (6) Barriers and Challenges of Data Manipulation, identifying the limitations and potential risks data manipulation can induce. Based on this codebook, we identified the key themes from the paper worksheets and interview data. Through axial coding and affinity diagramming, we then organized and synthesized the codes [15]. Two lead authors independently coded the data and addressed discrepancies through discussions. The themes were refined collaboratively, reaching consensus through extensive deliberation.



Data about me today	**Data Manipulation** using the final selected type
Work Hours 09:53 - 18:03 (7:10)	→ 11:53 - 18:03 (5:10)
Workout Duration Cardio: 28:51 Weight Training: 30:00	→ Cardio: 8:51 Weight Training: 30:00

⁉️ Briefly explain the reason for manipulating the data as shown above!

Work hours: Since I usually leave work after about 5 to 6 hours, so I thought reducing the recorded time by 2 hours might make things line up better.
Workout: It felt strange doing both cardio and weight training, giving them equal time, so I manipulated the data a bit.

Please describe today's self-tracking data manipulation experience in one sentence.
(ex. Increase: Exaggerating today workout data made me feel good, and made me think I should exercise harder tomorrow.)

By slightly cutting down the logged work hours, it felt like I stayed at work a reasonable amount before heading out. The same goes for cardio —it now feels like I just did a light wrap-up session.

How did manipulating yesterday's self-tracking data affect today's self-tracking?
(ex. Increase: Since I exaggerated the amount of exercise I did yesterday, I worked harder and exercised more today.)

By separating the types of workout and recording only the cardio part as slightly reduced, it seemed like the workout time hadn't changed much compared to usual, but I was able to achieve a sense of balance in the record.

Based on the today's self-tracking data manipulation, how will you plan for tomorrow?
(ex. Increase: Since I exaggerated today's workout time from 1 to 2 hours, I plan to exercise an extra one hour tomorrow.)

It seems like moderately reducing the data is more effective.
I will keep trimming about 1/3 from each part so that I end up with 30-50% more time left than originally planned.

Figure 4: Translated Worksheet example created by the participant. After choosing a type of data manipulation, the participants applied it repeatedly for a week. They recorded the original data and then manipulated according to the chosen manipulation type. They answered questions about the reasons, experiences, and the ongoing impact of data manipulation. The original worksheet are attached as Figure 6 in Appendix B.

4 FINDINGS

In this section, we delineate our findings into four themes: (1) our participants' perceptions and application of data manipulation, (2) the effects of data manipulation on self-reflection, (3) the impacts of data manipulation on behavior, and (4) the challenges and limitations associated with data manipulation.

4.1 General Orientation

No participant dropped out of the study. Since they were accustomed to diligently reflecting on the collected data without any adjustments, participants initially found the concept of manipulating the values and structure of their data unfamiliar. They initially hesitated and struggled with how to manipulate their own data, requiring some time to fully embrace the concept. However, as the exploratory workshops progressed, they began generating and refining increasingly diverse ideas, stimulated by their earlier ideas and researcher's facilitation of their engagement. In group setting, participants often drew inspiration from each other's ideas and attempting similar approaches with their own self-tracking data. Moreover, when participants tracked similar types of data and had comparable experiences, they deeply resonated with each other's ideas and built upon them to further develop their own thoughts. Through this process, participants came to view data manipulation as engaging and enjoyable.

During the one-week field trial, no participants chose to delete data. Five participants (P1, P2, P3, P5, P7) chose "insert", one participant (P10) chose "increase", one (P4) selected "reduce", and three (P6, P8, P9) opted to "transform" their data. During the experiment, three participants (P1, P5, P7) selected additional types, with P1 and P7 choosing "increase", and P5 opting for "transform" to enhance flexibility.

The exploratory workshops and the subsequent week-long field trial enabled participants to break free from their habitual reflection process and expand their perspectives, which was previously confined to the boundaries of their collected data. P5 noted that data manipulation gave him the enlightening experience: *"I used to think it was crucial to focus solely on numbers which is written on the self-tracking system. But through data manipulation, I felt liberated from the 'constraints imposed by the collected data.' I could think about the data more freely, asking myself, 'What might this mean?'"* Moreover, participants commented that data manipulation allowed them to develop a more meaningful and improved self-tracking practice, breaking out of the habitual and even automatic patterns of their self-tracking. The interactive process of manipulation provided them with the way to deeply reflect on what they recorded, describing this experience as worthy because they could move to the self-tracking that *"reflects my lifestyle"* (P6) and self-tracking that *"records my true activities"* (P4).

4.2 Approaches to Data Manipulation

Although five types of data manipulation—Insert, Delete, Transform, Increase, and Reduce—were provided, participants generated diverse ways to apply these techniques to their self-tracking data. Table 3 summarizes how participants applied each type of data manipulation, along with specific examples.

Diverse applications of data manipulations could be grouped into several approaches, as they were carried out with common goals or intentions. The five approaches and their illustrative examples of data manipulation are as follows.

1. Modify Broken Data Participants corrected inaccurate data values in their datasets to better align the recorded data with their self-perception and experience, creating a more reliable dataset for self-analysis. For example, P4, who tracks the sleep hours, modified his automatically recorded sleep time from 7 hours and 40 minutes to 5 hours and 40 minutes. He recalled being disturbed by strange noises during the night and felt that he would not be this tired if he had actually slept for 7 hours and 40 minutes.

2. Cleanse Data Participants managed their datasets by deleting specific entries, such as repetitive, out-of-context, or outlier data. This enhanced the dataset's clarity and objectivity. Some also removed data based on subjective criteria, such as entries they deemed unnecessary or undesirable. For example, P9 deleted emotionally charged journals that she found embarrassing or irrelevant, explaining that those memories felt unnecessary for reflection and distracted her from meaningful insights.

3. Impute Data Participants filled in missing data to address perceived gaps. This often involved adding plausible or arbitrary entries to acknowledge unrecorded activities, ensuring the dataset appeared complete and reflective of their efforts. For instance, P7, who runs for his workout, added a plausible entry based on their other physical activities that day, such as walking or light jogging, when he did not run.

4. Enrich Data Participants enriched their datasets by adding missing contexts or qualitative details, making the data more comprehensive and easier to interpret. For example, P5, who is tracking his spending, transformed the data by converting purchase amounts into different currencies to reflect emotional satisfaction: *"If I feel the purchase is worthwhile, I could make it look bigger by using larger denominations such as Venezuelan currency. If it feels like a waste, it will record it in smaller units like dollars or euros. It doesn't misrepresent the fact, but the way it looks does. I think it would help me see how much satisfaction I got from a purchase at a glance".*

5. Reinterpret and Adjust Data Participants revisited and adjusted past data from their current perspective, often tailoring the data to make it more intuitive or relevant to their current needs. For example, P4, who is tracking his work hours, excluded standard core hours (i.e., specific time period during the workday when employees are expected to be present and available for work) to identify the time when he voluntarily commits himself to focus. Also, P10, who writes journals, revisited old journal entries and revised negative events in a more neutral or positive light, aligning the records with how they now viewed those experiences.

4.3 Revitalize and Refresh the Self-reflection

4.3.1 Shaping Fresh and New Self-reflection. Freely manipulating data sparked various forms of self-reflection. Participants voiced that repeatedly encountering similar or even identical data and reflections eventually rendered the process of self-tracking "stagnant and meaningless." P4 stated, *"When I reflect on my workout data, it's always the same—same time, same goal, same routine. It's so repetitive that I don't even know how to interpret it anymore. It just feels stagnant, like it's not telling me anything useful."* They expressed a desire for more expansive and awakening self-reflection. By breaking free from the constraints of their previous data interpretations, participants were able to reinterpret the data from fresh perspectives and uncover new meanings. For example, P10, who toned down the emotions in the journal he wrote, expressed the feeling that, *"Now, when I look back at this, there are some emotions I can feel right now. But if all the emotions are written down as they were, I just follow them again. By reducing them, I actually feel richer*

Type (Total Freq.)	Application (Freq.)	Example
Insert (52)	Add Missing Data (10)	Add workout data that was not collected because the wearable tracker was not worn
	Contextualize Data (20)	Add "leg pain" as a factor to step count data
	Add Imagination of Future (6)	Add that I published a book as an author in the future in my journal
	Arbitrarily Add False Data (6)	Falsely add workout data I did not actually do yesterday without basis
	Falsely Add Missing Data with Reasoning (10)	Given that I walked a lot, assume that I ran some of the distance
Delete (12)	Remove Outlier Data (2)	Remove large, irregular expenses from the finance data
	Remove Redundant Data (2)	Remove repeated words or sentences from a journal
	Remove Irrelevant Data (2)	Delete words or sentences from a journal that do not make sense in the overall context
	Remove Data Containing or Evoking Negative Emotions (3)	Delete short-distance driving data because it triggers guilt about environmental impact
	Remove Meaningless Data (3)	Remove the "heart rate" factor from workout data because it is something I cannot control
Transform (45)	Transform Data for Better Understanding (19)	Convert the price of a luxury item I want to buy into "the number of workdays needed to afford it"
	Transform Data to Reflect Personal Thoughts and Feelings (17)	Adjust the currency to larger units for satisfying purchases and smaller for unsatisfying ones
	Reframe Emotions in Data (4)	Rewrite a negative journal positively
	Reconstruct Past Data from a Current Perspective (5)	Revise an old journal to reflect what I have learned since then
Increase (24)	Increase Inaccurate Data (6)	Increase weight data that was recorded inaccurately
	Specify and Detail Data (8)	Break down vague to-do list data like "Take care of the cat" into specific tasks (e.g., feed, give medication, clean the litter box)
	Slightly Increase Data (10)	Record 50 push-ups, though I only managed 45
Reduce (23)	Decrease Inaccurate Data (4)	Reduce inaccurately recorded data that says I watch more YouTube than I actually do
	Decrease Data According to Standard Criteria (10)	Subtract core working hours (i.e., 5 hours) from my total working hours data
	Remove Emotion from Data (2)	Tone down emotions in a journal that is filled with strong emotional content
	Slightly Decrease Data (7)	Record that I used smartphone for 8 hours when I actually used it for 9 hours

Table 3: The table exhibits application and its examples by each type of data manipulation. Participants applied data manipulation in various ways, and the applications are categorized into five strategies based on common objectives or purposes.

emotions. It feels like I can interpret it in more diverse ways." Participants leveraged several data manipulation approaches to gain new and meaningful reflections.

One effective approach to manipulating the data for fresh self-reflections was to reshape previous data into more personally resonant terms. It created more powerful and intuitive self-reflections. For example, there are cases of converting spending amounts into the equivalent number of school meals (P8), transforming the time wasted on Instagram into a list of what they could have done during that time (P3). It offered a fresh lens to view their actions and confront aspects of themselves they just overlooked: *"When I see that I spent 50,000 or 60,000 KRW in a day, it doesn't seem like a small* *amount, but I just accept it. But when I think of it as 10 school meals–5 days' worth –it really hits home. The seriousness of spending what I'd eat for 5 days in one day."* (P8) It encouraged the participant to develop constructive plans for improvement. By manipulating their data, they opened up new opportunities for self-reflection, creating a pathway for more intentional and actionable insights.

Another approach leading to fresh reflections was manipulating data to generate various future scenarios and possibilities. This encouraged the participants to reflect on their future selves, extending their focus beyond the present. One participant (P10) remarked that he could imagine not only realistic future scenarios but also extreme or unrealistic ones by manipulating the past data, allowing him to envision a richer and more proactive future. Another participant shared that the reflection on the future even led to create a concrete plan in the present: *"Diaries are usually about past events. But when I rewrite the journal entries as I want, future stories naturally come into the diary, and those commitments to the future unconsciously turn into plans"* (P9). By reflecting on their future selves, participants became more aware of their expectations and the version of themselves they aspired to be. While participants found it difficult to engage in self-reflection about their ideals and the person they wanted to become, data manipulation provided valuable insight into these aspirations.

In addition to specific strategies, the act of data manipulation itself appeared to encourage new reflections. As participants should decide what and how to manipulate data during the manipulation process, they began to assess and evaluate the data more critically. This pushed the participants to go beyond a previous understanding of the data, expanding the way they viewed and reflected on the data: *"Manipulating data means there's something missing or unsatisfactory. So naturally, checking the data's quality becomes necessary to find out what's lacking."* (P1). Assessing whether the data was satisfying or disappointing to themselves enabled self-anchoring, providing insights like "I'm doing this much" and "I think I am lacking." Self-anchoring, as a way of understanding one's self-concept (i.e., "Who am I?") [69], enriched participants' self-reflection and self-knowledge through the data manipulation process.

4.3.2 From Rumination toward Reflection. Many participants found that data manipulation helped break the cycle of rumination and proceed with constructive self-reflection. When seeing shortfalls in their performance metrics such as fewer steps walked, fewer hours of productive work, or fewer calories burned, participants often felt despair, guilt, or a sense of personal failure. Some participants blamed themselves for not living up to the standards they had set. Over time, these feelings eroded their motivation, preventing meaningful reflection. By reflecting on the reasons behind their inability to achieve their goals while manipulating their data, participants could reinterpret shortfalls as understandable exceptions rather than as signs of their intrinsic inabilities. This reinterpretation allowed them to step back from rumination and create a more emotionally safe space for self-reflection. In this space, they could gently explore why things did not go as planned, which allowed for constructive reflection. Ultimately, this shift gave motivation to continue their self-tracking routines in the following days: *"When I log that I exercised even though I didn't, it actually reduces my guilt. Not being able to record a workout feels stressful—it's like, "I didn't exercise today. Am I just weak-willed? What's wrong with me?" It brings a lot of self-criticism and guilt. By simply manipulating the record to show I exercised, I feel like it would make it easier for me to work out next time."* (P1).

Shame also emerged as a strong emotional response to self-tracking data. Some participants felt uncomfortable or embarrassed by certain entries that reminded them of habits they found distressing or emotional events they had recorded. By altering or removing these entries, participants created an emotional space that allowed for more positive self-reflection. Even though many participants were the only ones with access to their records, they found relief by deleting or manipulating these entries. As P9 expressed, *"I want to get rid of things that are embarrassing to look back on. Even though I am the only one who sees my journal, some memories are ones I'd rather erase from my life."* Nevertheless, many participants felt a significant barrier when it came to the idea of deleting data. Participants regarded complete deletion as a destructive act that made the data entirely inaccessible. Many of them, even when the data triggered negative emotions, were reluctant to remove it because they perceived the data—both positive and negative—as reflecting themselves and a part of their personal assets, resonating the belief that data is an important personal collection or a facet of their life [19]: *"Deleting the data [that triggered unpleasant emotions] makes me feel a bit safer, but at the same time, I want to bring it back. After all, these are my thoughts, so it is too valuable to just discard."* (P10) There was a conflicting feeling of wanting to delete the embarrassing data, yet still wanting to keep it. Therefore, rather than deleting the data, they expressed a preference for hiding it in a backlog or making it less easily accessible.

4.3.3 Refining Self-reflection. Data manipulation provided participants with opportunities to refine their self-reflection process. By correcting inaccuracies and filling in gaps between collected and their perception, participants were able to create more complete datasets that supported a better understanding of their experiences. Previously, many participants have experienced uncertainty or confusion during self-reflection, especially when their lived experiences did not align with the data they had collected. This mismatch often created discomfort, making it challenging to engage in meaningful reflection. However, when participants adjusted the data to better match their perceptions, they were able to restore a sense of coherence, reducing some of the anxiety they felt. This process helped them refine their existing reflections. For example, P9, who added more details to hastily written journal entries, expressed: *"I used to think I wasn't interested in people. But after I rewrote my journal, it turns out I'm more interested than I thought."*

Participants also mentioned feeling uncertain when repetitive or outlier data made it difficult to discern meaningful patterns. In some cases, unusual data points, such as those arising from exceptional circumstances, interfered with identifying clear trends. For example, P2, who had recently moved, found that one-time moving expenses distorted her spending patterns. This made it difficult to understand her usual living costs, leaving her feeling uncertain. After removing the moving-related expenses, P2 could get clearer picture of regular consumption habits. By manipulating the data, participants could alleviate some confusion and go through more clear self-reflection.

4.4 Psychological Boost for Behavioral Momentum

4.4.1 Positive Illusions. At the exploratory workshop, several participants mentioned that they were unsure about the effectiveness of manipulating data as they want, doubting whether it could truly alter their memories. P1 stated, *"Since I know I didn't exercise, I'm not sure how effective it would be to manipulate that I did"*. This skepticism highlighted an initial uncertainty about whether data manipulation could have any meaningful impact on their behavior. However, after engaging in the one-week field trial, many participants discovered that data manipulation indeed served as a powerful catalyst, fostering a positive mindset toward behavioral engagement and change. Although the manipulated data did not alter the actual memories, the participants experienced a psychological illusion by believing in the manipulated data as if it were real: *"Even though I knew I did 55 reps of push-ups, the manipulated data clearly shows 60, so I felt like 'I must have done 60 on that day,' and it felt like I was deceiving myself. It was like experiencing such a placebo effect"* (P1). This illusion led them to tell themselves 'I did this much last time, so I can do more this time'. Such a sense of active control, which is similar to the concept of autosuggestion [51], boosted their confidence but also helped increase their exercise levels and reach higher goals. Beyond exercise tracking, similar effects were observed in other domains. For instance, P9, who wrote an emotional diary after a test failure, rewrote the narrative from a more rational perspective. This process of manipulation allowed P9 to view the failure as less significant, subsequently increasing self-belief and motivation to continue pursuing their activities.

4.4.2 Sustaining Momentum. The act of data manipulation transformed into a sense of responsibility, providing the motivation to persist in pursuing their goal-oriented behaviors. Although data manipulation effectively mitigated the negative emotions generated by self-tracking records and helped regain motivation to continue acting, it did not completely eliminate the inherent sense of guilt associated with manipulating one's data. Participants characterized this lingering guilt as *"a feeling of being indebted."* This underlying guilt eventually evolved into self-accountability. It encouraged participants to achieve their goals without relying on further manipulations. For others, like P1, the act of data manipulation grew into something more than simply altering records—it became a form of *"keeping a promise."* This process provided the momentum to uphold the commitments and maintain active behavioral engagement.

4.5 Lesson Learned from the Study

4.5.1 Importance of Balanced Manipulation. Participants encountered challenges in determining the appropriate extent of data manipulation. They found out that manipulating data too extensively could distort both perception of the recorded data and oneself. One illustrative episode comes from P4, who initially deducted 5 hours (i.e. core hours) from his work time in an attempt to gauge how much time he spent voluntarily focusing outside of core hours. However, subtracting 5 hours proved excessive because the remaining work hours appeared too few after the deduction. This created a distortion in his self-perception. He saw himself as having worked very little, felt a lack of responsibility, and even labeled himself as lazy. In response, P4 experienced an over-ambition, pushing himself harder in an effort to compensate by logging even longer hours. Although he eventually recorded 10 hours of work, thereby still maintaining a substantial total even after the 5-hour deduction, his body and mind became overstrained. It eventually disrupted his self-tracking practices. These distortions led to unrealistic goal-setting and poor decision-making. To address these unintended consequences, P4 recalibrated his approach by experimenting with different adjustments over several days. Ultimately, he reduced his work hours by only 2 hours instead of 5, resulting in a more realistic and sustainable record. P4's self-experimentation culminated in a balanced approach that left him with the positive impression: *"It feels like I've spent just the right amount of time at work."* In addition, participants emphasized the importance of periodically confronting reality and staying connected to their authentic selves to avoid misperceptions that manipulated data could induce. While working to find a range of adjustments that wouldn't diverge too far from their true selves, they discovered that revisiting the original data and reflecting on their actual habits was crucial for preventing the unintended consequences of excessive manipulation.

Similarly, other participants demonstrated that carefully moderated data manipulation can yield more positive outcomes. Many of these individuals opted to limit the degree of their manipulation by manipulating data only within an "understandable and realistic range." This meant making adjustments only in situations where they typically could have met their goals but fell short due to external factors (e.g., *"when it was too hot to run and I couldn't meet my target"* (P7) or *"when I couldn't meet my activity goal due to not feeling well"* (P2)). Participants explained that manipulating data outside of this realistic range tended to evoke feelings of guilt and frustration, further widening the discrepancy between their actual performance and the modified data. They avoided merely tweaking numbers to create the appearance of progress. Instead, they used manipulation as a strategy to provide genuine self-insight and a meaningful psychological boost.

4.5.2 Not a Universal Cure. Not all self-tracking users are amenable to the same data manipulation type. Participants showed clear preferences toward specific types among five typology based on their unique goals, values, and needs. For instance, those seeking positive reinforcement and behavior change found the "Increase" the most meaningful and applicable. In contrast, participants who desired fresh insights from an alternative perspective valued the "Transform", while P9—who maintains a fact-based journal—found "Reduce" incompatible with her data, stating that reducing factual entries felt challenging because it forced her to decide what to omit. However, P10, who typically records more emotion-driven entries rather than factual accounts, found the "Reduce" an engaging approach that helped alleviate overwhelming emotions. Even among self-trackers who keep journals, individual preferences varied greatly. Similarly, the applicability of various data manipulation methods depended on the domain of self-tracking data. Not all types of data can be manipulated with equal ease. Especially, in the context of financial self-tracking where accuracy is paramount, adjusting data value offers no benefit because understanding the true state of income and expenses is essential for decision-making. Therefore, altering numerical values such as Increase, Reduce could not be adopted to the financial self-tracking data. P8 stated, *"In financial self-tracking, I think it's crucial to stay grounded in facts because manipulating monetary values could cause real harm. On the other hand, data like records about workout and car driving are personal and don't affect anyone else, so tweaking those seems acceptable, but financial data feels different—it needs to remain accurate."* In a similar vein, P5 agreed with the potential of manipulating the personal data for self-motivation, but emphasized the importance of maintaining the accuracy of financial data because it represents relations with others: *"It represents transactions with others, so altering it could compromise its integrity as data reflecting monetary relationships."* Ultimately, these varied responses highlight that data manipulation is not a universal remedy. Its effect and acceptability depends not only on individual preferences but also on the nature of the data being tracked.

5 DISCUSSION AND DESIGN IMPLICATIONS

In this study, we applied an alternative approach of data manipulation in self-tracking with a focus of self-reflection. It was an exploratory study to investigate how self-trackers apply data manipulation, and what experiences as well as perceived effect they got throughout the process. In this section, we reflect on the role of data manipulation, highlighting its dual nature. We then propose several design implications for integrating data manipulation into future PI systems.

5.1 Reflections: The Dual Nature of Data Manipulation

This study applied the data manipulation to individuals who were stuck in the process of self-reflection, exploring its effects and experiences. Manipulating data helped self-trackers gain awareness of alternative perspectives and develop new insights, opening possibilities to move to higher levels of reflection [28]. While repeated or similar stimuli often lead individuals to fixate the thought patterns [35], data manipulation encouraged users to explore and interact with their data in diverse ways. Data manipulation diversified how users explore their data and engage in reflective thinking. It showed the potential to achieve the concept suggested by Rapp and Tirassa, *user-controlled proliferation of takes on the user's data* [58]. Additionally, the positive illusions created by manipulating data could promote positive behavioral changes. These illusions function similarly to mental imagery manipulation [31], where envisioning better outcomes enhances positive emotions, self-confidence, and self-belief, ultimately improving performance [8]. Lastly, data manipulation played a role in alleviating negative emotions, which supported sustained self-reflection and provided the possibility of positively influencing the maintenance of actions. It could act as a psychological buffer [73], helping participants avoid the demotivating effects of negative emotions. This alleviation of negativity could create a positive feedback loop, providing momentum to maintain existing behaviors [33, 45, 63]. Overall, the approach of data manipulation aligns with Pantzar and Ruckenstein's perspective of "situated objectivity" [56]. This perspective suggests that the mechanical objectivity of data does not hold inherent value. Instead, subjectively interpreting data and assigning personal meaning to it enhances its value, leading to deeper self-reflection and meaningful change in their lives. Data manipulation also aligns with the concept of flexibility in self-tracking. In PI ecosystem that supports flexibility individuals can configure data in preferred format and present data in personalized ways that reflect their evolving interests and needs. This enables self-trackers to derive personally meaningful insights, ultimately facilitating deeper self-reflection [1, 5, 6, 36]. By allowing individuals to explore and engage with their data in adaptable ways, data manipulation parallels to the flexibility. Moreover, it may also extend beyond flexibility: by allowing self-trackers to actively reconstruct the values or structure of existing data, data manipulation could prompt them to reconsider prior interpretations and support fresh and enriched self-reflection.

Nevertheless, it is revealed that data manipulation does not always produce positive outcomes. When overused, data manipulation distorted both the data and one's self-perception. It could mask one's true condition and self-awareness [11, 49]. This can prevent meaningful growth and exacerbate underlying issues that remain unaddressed. Moreover, the same approach to data manipulation could not be universally applied to the circumstances of various self-trackers and their data domain. As self-trackers adjust various self-tracking tools and styles based on their objectives and everyday circumstances [18, 60], it would be beneficial to tailor data manipulation to their specific contexts, needs, and goals. In conclusion, it implies that data manipulation has a dual nature: if used excessively and indiscriminately, it may lead to distortions and unintended negative outcomes, yet when applied judiciously, it serves as a valuable tool for self-reflection and behavioral change. By balancing its mindful limits with the potential benefits, data manipulation could serve as a valuable addition to future PI systems.

5.2 Design Implications for Integrating Data Manipulation into PI Systems

Self-trackers can engage in the self-tracking process more accessibly and ubiquitously through PI systems leveraging digital technologies [57] (e.g., mobile fitness apps, fitness trackers). If data manipulation is effectively integrated into these digital PI systems, users could experience significant opportunities for self-reflection and subsequent actions. Therefore, we propose design implications for how designers can effectively integrate data manipulation into digital PI systems.

5.2.1 Lowering the Barrier to Data Manipulation. There are certain user groups that could potentially benefit greatly from leveraging data manipulation (e.g., those who were stuck during self-reflection), even though not every self-tracker may be suited for data manipulation. However, as participants initially hesitated and found it challenging to manipulate data values and structures (see Section 4.1), the approach to data manipulation could be a barrier. Since PI data is often considered valuable for its objectivity and integrity [54], making the idea of manipulating data would be unfamiliar to many self-trackers. Despite the potential benefits data manipulation could offer, they could pose the question of "Is it okay to change my data?", creating a barrier to trying it out.

Therefore, it is important to introduce the new opportunities and benefits data manipulation can provide, in order to encourage users to explore it more openly. To detect and introduce user the concept of data manipulation, PI system could track not only traditional metrics (e.g., steps, heart rate) but contextual factors such as emotional state or overall progress. When the self-tracker appears to be in a stagnant phase or show signs of demotivation, the system could suggest the possibility of data manipulation with a simple prompt like "Need some fresh insights? How about reshaping your data in more creative ways?" Once users are introduced to the idea, providing opportunities for simulating data manipulation can be helpful for exploring its potential. To do so, sandbox mode (i.e., a mode where users get freedom to explore and experiment, commonly used in games [67]) could be an option. Allowing self-trackers to experiment data manipulation directly on their small-scale datasets in a non-permanent way can help them experience the possible outcomes and discover benefits. This sandbox mode would help reduce the barriers they feel towards data manipulation.

5.2.2 Guide to Choosing the Best Data Manipulation Type. Depending on an individual's needs, goals, and values, as well as the specific area of self-tracking data, there are preferred types of data manipulation. Participants in our study faced challenges or resistance when confronted with types of data manipulation that were difficult to apply to their data or that could lead to negative outcomes (see Section 4.5.2). To reduce this confusion, it would be helpful to recommend the type of data manipulation most suited to a self-tracker's style and the specific domain of data they are tracking. Therefore, before performing data manipulation, it would be useful to ask self-trackers a series of questions about their self-tracking data and their values around self-tracking (e.g., (1) What type of data are you collecting? (2) Why are you engaging in self-tracking? (3) What do you hope to achieve with self-tracking?). Based on this information, PI systems could go through a user profiling process. The user group could be organized based on frameworks like Rooksby et al.'s 'Five Styles of Personal Tracking' (directive tracking, documentary tracking, diagnostic tracking, collecting rewards, and fetishized tracking) [60]. Based on the profile, the most appropriate type of data manipulation could be recommended. For instance, a fetishized tracker who enjoys collecting and interacting with data might be encouraged to explore the "Transform" type of manipulation to gain fresh and new perspectives.

In addition to considering a self-tracker's style, it is also important to recommend the most suitable type of manipulation based on the domain of data being tracked. For instance, the domain such as financial data requires accurate tracking, so the type of manipulation that does not alter data values was found to be more suitable. Therefore, PI systems should carefully consider the characteristics of self-tracking domain and the associated goals and needs when recommending the most appropriate way of manipulation. Beyond offering recommendation, it might also be necessary to clearly define explicit boundaries, indicating which manipulations should or should not be permitted. In real-world scenarios, users are likely to engage with a broader range of self-tracking data domains than those examined in this study. This could increase the risk of unintended negative consequences from certain types of data manipulation. For example, some users could track data closely related to their personal safety and health, such as chronic diseases [47] or heart disease [44]. In these contexts, accurate and precise data is crucial for continuous tracking and health management. Users should thus be supported in enhancing data accuracy and minimizing discrepancies between their actual behaviors and the recorded data, rather than being offered new or entertaining perspectives. In such critical contexts, PI systems should establish the guidelines specifying permissible data manipulations and firmly guide users to these boundaries.

5.2.3 Supporting the Application of Data Manipulation. Even when a user selects a data manipulation type, there are various ways it can be applied. Participants in our study found themselves contemplating and generating various ideas on how to apply a chosen data manipulation type to their own data (see Section 4.2). Since users may feel uncertain about how to apply a particular manipulation type to their data, it would be helpful to provide support for this application process. One directive approach could be for the PI systems to gather information about the user's intentions and desired outcomes and then directly manipulate their self-tracking data. However, system-driven features have the limitation of failing to capture the user's nuanced needs and preferences [20], so it would be preferable for users to take control and lead the data manipulation process. To effectively assist users while respecting their agency, integrating an assistant powered by Large Language Models (LLMs) could be a valuable solution. LLMs can interpret the user's needs and provide responses that are tailored to their specific situation, since they are capable of understanding the user's context and nuances [71], During conversations with users, the LLM-based assistant could offer tailored suggestion and guidance on how to apply data manipulation. Moreover, LLM-based assistant could generate reflective questions [39] to assist with the users to consider different ways to manipulate their data. For example, the assistant could offer an question like, "What would happen if you increase the workout time by 30 minutes yesterday evening?"

Additionally, one of the key considerations in the application of data manipulation is ensuring an "appropriate scope" of manipulation, as excessive data manipulation led to negative outcomes. To help users reflect on whether their manipulation is within an appropriate range, it would be beneficial to allow them to confront the reality by viewing the original data and maintaining a connection to their authentic selves. To do so, we could think about the comparison feature to visually compare the data before and after manipulation. This could be implemented through a button or slider that enables users to slide between two datasets, displaying them side by side. This feature could help users to reflect on the gap between the original and manipulated data and prompt them to ask themselves, "Is this manipulation too drastic?" Alternatively, PI systems could support users in conducting iterative self-experimentation, as P4's experience in our study. While users apply different levels of manipulation, PI systems could prompt them to reflect on their overall progress, and positive changes, and emotions resulting from each level of manipulation. This iterative process of reflecting on oneself and providing self-feedback on each level of manipulation could help them fine-tune the balance of their data manipulation over time.

6 LIMITATIONS AND FUTURE WORK

Our study presents some limitations and areas for future research. First, the small sample size of 10 participants may have constrained the diversity of interpretations, applications, and experiences related to data manipulation. A larger participant pool would allow for a broader range of insights into how users engage with data manipulation and its effects. Another limitation was the narrow range of self-tracking data domains represented by the participants. Although efforts were made to recruit participants with diverse self-tracking practices, most of the data focused on domains such as workouts, finances, productivity, and journaling. To explore the effects of data manipulation across a broader spectrum of self-tracking practices, future studies should include participants from a wider variety of fields.

Additionally, the one-week period for applying data manipulation seemed to be too short to fully explore its potential impact. While exploratory workshops helped the participants in generating and specifying ideas for data manipulation, participants noted that translating these ideas into the field trial required significant time for fine-tuning. Adjusting data manipulations to fit their individual contexts and determining the optimal scope for supporting their self-tracking involved trial and error. Thus, it would be beneficial to allow self-trackers to engage with data manipulation over a longer period. However, it remains unclear whether users would consistently maintain engagement with data manipulation practices in real-world scenarios if the duration were extended. Therefore, future studies should explore a longer application period to examine various aspects more thoroughly, such as the ongoing effects of data manipulation and how user engagement evolves over time.

A final limitation is that data manipulations took place on paper worksheets rather than within participants' actual PI tools, which restricted us to observe their downstream effects. While the paper-based approach facilitated flexible expression, active engagement, and side-by-side comparison of original and manipulated data, it inherently limited the possibility to examine how such manipulations would be subsequently represented in real systems (e.g., through visualizations or summaries). As a result, we were limited in exploring how the integration of manipulated data with other data and its representations might further influence participants' self-reflection, actions, and experiences. To better understand these downstream effects beyond the act of manipulation itself, future work could involve developing a PI system that incorporates data manipulation and deploying it in self-tracking contexts.

7 CONCLUSION

In this paper, we introduce data manipulation as an alternative approach. Our goal was to understand how self-trackers perceive and apply data manipulation, and the experience and impact on self-reflection and subsequent actions. Our findings from the exploratory workshops and one-week field trial reveal that self-trackers could utilize data manipulation across the broad spectrum based on their personal contexts and needs. Data manipulation gave opportunity to revitalize and renew self-reflection, while also boosting motivation for behavioral change and engagement. However, the approach also carries distorted perceptions and limitations, highlighting the importance of mindful use on manipulation practices. We conclude with design implications about PI systems with data manipulation, guidelines for applying data manipulation and selecting the appropriate scope. We hope our work paves the way for future research in personal informatics using data manipulation as an approach.

ACKNOWLEDGMENTS

We thank our study participants for their contributions. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (RS-2024-00458557) and the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2022-0-00064).

REFERENCES

[1] Parastoo Abtahi, Victoria Ding, Anna C Yang, Tommy Bruzzese, Alyssa B Romanos, Elizabeth L Murnane, Sean Follmer, and James A Landay. 2020. Understanding physical practices and the role of technology in manual self-tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–24.

[2] Deemah Alqahtani, Caroline Jay, and Markel Vigo. 2020. The Role of Uncertainty as a Facilitator to Reflection in Self-Tracking. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) *(DIS '20)*. Association for Computing Machinery, New York, NY, USA, 1807–1818. https://doi.org/10.1145/3357236.3395448

[3] Jessica S Ancker, Holly O Witteman, Baria Hafeez, Thierry Provencher, Mary Van de Graaf, and Esther Wei. 2015. "You get reminded you're a sick person": personal data tracking and patients with multiple chronic conditions. *Journal of medical Internet research* 17, 8 (2015), e202.

[4] Chris Argyris. 2002. Double-loop learning, teaching, and research. *Academy of management learning & education* 1, 2 (2002), 206–218.

[5] Amira Ayobi, Paul Marshall, and Anna L. Cox. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Selfcare. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA). Paper presentation.

[6] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L Cox. 2018. Flexible and mindful self-tracking: Design implications from paper bullet journals. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–14.

[7] 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 '21)*. Association for Computing Machinery, New York, NY, USA, Article 246, 12 pages. https://doi.org/10.1145/3411764.3445505

[8] Simon E Blackwell, Michael Browning, Andrew Mathews, Arnaud Pictet, James Welch, Jim Davies, Peter Watson, John R Geddes, and Emily A Holmes. 2015. Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: a randomized controlled trial. *Clinical Psychological Science* 3, 1 (2015), 91–111.

[9] Tali Boritz, Ryan Barnhart, Lynne Angus, and Michael J Constantino. 2017. Narrative flexibility in brief psychotherapy for depression. *Psychotherapy Research* 27, 6 (2017), 666–676.

[10] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[11] Zoë Chance, Michael I Norton, Francesca Gino, and Dan Ariely. 2011. Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences* 108, supplement_3 (2011), 15655–15659.

[12] 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 *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 142, 23 pages. https://doi.org/10.1145/3491102.3501991

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

[14] Eun Kyoung Choe, Naeun 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*. 1143–1152.

[15] Juliet M Corbin and Anselm Strauss. 1990. Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology* 13, 1 (1990), 3–21.

[16] 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.

[17] Elizabeth V. Eikey and Madhu C. Reddy. 2017. "It's Definitely Been a Journey": A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 642–654. https://doi.org/10.1145/3025453.3025591

[18] Chris Elsden, Abigail C Durrant, David Chatting, and David S Kirk. 2017. Designing documentary informatics. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. 649–661.

[19] 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.

[20] 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 '14)*. Association for Computing Machinery, New York, NY, USA, 667–676. https://doi.org/10.1145/2598510.2598558

[21] Daniel A. Epstein, Alan Borning, and James Fogarty. 2013. Fine-grained sharing of sensed physical activity: a value sensitive approach. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Zurich, Switzerland) *(UbiComp '13)*. Association for Computing Machinery, New York, NY, USA, 489–498. https://doi.org/10.1145/2493432.2493433

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

[23] Daniel A Epstein, Monica Caraway, Chuck Johnston, An Ping, James Fogarty, and Sean A Munson. 2016. Beyond abandonment to next steps: understanding and designing for life after personal informatics tool use. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 1109–1113.

[24] 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*. 829–840.

[25] Daniel A. Epstein, Alexandra 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*. 731–742.

[26] Clayton Feustel, Shyamak Aggarwal, Bongshin Lee, and Lauren Wilcox. 2018. People Like Me: Designing for Reflection on Aggregate Cohort Data in Personal Informatics Systems. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 107 (Sept. 2018), 21 pages. https://doi.org/10.1145/3264917

[27] 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*. 216–223.

[28] 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* (Brisbane, Australia) *(OZCHI '10)*. Association for Computing Machinery, New York, NY, USA, 216–223. https://doi.org/10.1145/1952222.1952269

[29] Salvador García, Julián Luengo, and Francisco Herrera. 2015. *Data Preprocessing in Data Mining*. Vol. 72. Springer International Publishing. 59–139 pages.

[30] 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 '15)*. Association for Computing Machinery, New York, NY, USA, 1305–1316. https://doi.org/10.1145/2750858.2804290

[31] Natalie Gukasyan and Sandeep M Nayak. 2022. Psychedelics, placebo effects, and set and setting: Insights from common factors theory of psychotherapy. *Transcultural psychiatry* 59, 5 (2022), 652–664.

[32] Cheng-Kang Hsieh, Hongsuda Tangmunarunkit, Faisal Alquaddoomi, John Jenkins, Jinha Kang, Cameron Ketcham, Brent Longstaff, Joshua Selsky, Betta Dawson, Dallas Swendeman, Deborah Estrin, and Nithya Ramanathan. 2013. Lifestreams: a modular sense-making toolset for identifying important patterns from everyday life. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems* (Roma, Italy) *(SenSys '13)*. Association for Computing Machinery, New York, NY, USA, Article 5, 13 pages. https://doi.org/10.1145/2517351.2517368

[33] Gyuwon Jung, Jio Oh, Youjin Jung, Juho Sun, Ha-Kyung Kong, and Uichin Lee. 2021. "Good Enough!": Flexible Goal Achievement with Margin-based Outcome Evaluation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[34] Elisabeth T Kersten-van Dijk, Joyce HDM Westerink, Femke Beute, and Wijnand A IJsselsteijn. 2017. Personal informatics, self-insight, and behavior change: a critical review of current literature. *Human–Computer Interaction* 32, 5-6 (2017), 268–296.

[35] Joe Khatena. 1970. Repeated presentation of stimuli and production of original responses. *Perceptual and Motor Skills* 30, 1 (1970), 91–94.

[36] Young-Ho Kim, Jae Ho Jeon, Bongshin Lee, Eun Kyoung Choe, and Jinwook Seo. 2017. OmniTrack: A Flexible Self-Tracking Approach Leveraging Semi-Automated Tracking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 67 (Sept. 2017), 28 pages. https://doi.org/10.1145/3130930

[37] Bran Knowles, Alison Smith-Renner, Forough Poursabzi-Sangdeh, Di Lu, and Halimat Alabi. 2018. Uncertainty in current and future health wearables. *Commun. ACM* 61, 12 (Nov. 2018), 62–67. https://doi.org/10.1145/3199201

[38] Rafal Kocielnik, Daniel Avrahami, Jennifer Marlow, Di Lu, and Gary Hsieh. 2018. Designing for Workplace Reflection: A Chat and Voice-Based Conversational Agent. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 881–894. https://doi.org/10.1145/3196709.3196784

[39] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection companion: a conversational system for engaging users in reflection on physical activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–26.

[40] Dorthe Brogård Kristensen and Minna Ruckenstein. 2018. Co-evolving with self-tracking technologies. *New Media & Society* 20, 10 (2018), 3624–3640.

[41] 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*. 557–566.

[42] Catherine Y. Lim, Andrew B.L. Berry, Andrea L. Hartzler, Tad Hirsch, David S. Carrell, Zoë A. Bermet, and James D. Ralston. 2019. Facilitating Self-reflection about Values and Self-care Among Individuals with Chronic Conditions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300885

[43] Meagan B. Loerakker, Jasmin Niess, Marit Bentvelzen, and Paweł W. Woźniak. 2024. Designing Data Visualisations for Self-Compassion in Personal Informatics. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 169 (Jan. 2024), 22 pages. https://doi.org/10.1145/3631448

[44] Stine Lomborg, Henriette Langstrup, and Thor O. Andersen. 2020. Interpretation as Luxury: Heart Patients Living with Data Doubt, Hope, and Anxiety. *Big Data & Society* 7, 1 (2020). https://doi.org/10.1177/2053951720924436

[45] Stine Lomborg, Nanna Bonde Thylstrup, and Julie Schwartz. 2018. The temporal flows of self-tracking: Checking in, moving on, staying hooked. *New Media & Society* 20, 12 (2018), 4590–4607.

[46] Gabriela Marcu, Anind K. Dey, and Sara Kiesler. 2014. Designing for collaborative reflection. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare* (Oldenburg, Germany) *(PervasiveHealth '14)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), Brussels, BEL, 9–16. https://doi.org/10.4108/icst.pervasivehealth.2014.254987

[47] Mollie McKillop, Lena Mamykina, and Noémie Elhadad. 2018. Designing in the dark: eliciting self-tracking dimensions for understanding enigmatic disease. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–15.

[48] Jack Mezirow. 2018. Transformative learning theory. In *Contemporary theories of learning*. Routledge, 114–128.

[49] Franklin G Miller, David Wendler, and Leora C Swartzman. 2005. Deception in research on the placebo effect. *PLoS medicine* 2, 9 (2005), e262.

[50] Adrián Montesano, João Tiago Oliveira, and Miguel M Gonçalves. 2017. How do self-narratives change during psychotherapy? A review of innovative moments research. *Journal of Systemic Therapies* 36, 3 (2017), 81–96.

[51] Kasia A Myga, Esther Kuehn, and Elena Azanon. 2022. Autosuggestion: a cognitive process that empowers your brain? *Experimental Brain Research* 240, 2 (2022), 381–394.

[52] 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 '20)*. Association for Computing Machinery, New York, NY, USA, Article 6, 11 pages. https://doi.org/10.1145/3379503.3405662

[53] 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* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173745

[54] Mari Niva. 2017. Online weight-loss services and a calculative practice of slimming. *Health:* 21, 4 (2017), 409–424.

[55] Rikke Ørngreen and Karin Tweddell Levinsen. 2017. Workshops as a research methodology. *Electronic Journal of E-learning* 15, 1 (2017), 70–81.

[56] Mika Pantzar and Minna Ruckenstein. 2017. Living the metrics: Self-tracking and situated objectivity. *Digital health* 3 (2017), 2055207617712590.

[57] Amon Rapp and Federica Cena. 2015. Affordances for self-tracking wearable devices. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers* (Osaka, Japan) *(ISWC '15)*. Association for Computing Machinery, New York, NY, USA, 141–142. https://doi.org/10.1145/2802083.2802090

[58] 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.

[59] James P Robson Jr and Meredith Troutman-Jordan. 2014. A concept analysis of cognitive reframing. *Journal of Theory Construction & Testing* 18, 2 (2014).

[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*. 1163–1172.

[61] Herman Saksono, Carmen Castaneda-Sceppa, Jessica Hoffman, Magy Seif El-Nasr, Vivien Morris, and Andrea G. Parker. 2018. Family Health Promotion in Low-SES Neighborhoods: A Two-Month Study of Wearable Activity Tracking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173883

[62] Moushumi Sharmin, Andrew Raij, David Epstien, Inbal Nahum-Shani, J. Gayle Beck, Sudip Vhaduri, Kenzie Preston, and Santosh Kumar. 2015. Visualization of time-series sensor data to inform the design of just-in-time adaptive stress interventions. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 505–516. https://doi.org/10.1145/2750858.2807537

[63] Sehee Son and Youn-Kyung Lim. 2023. Balancing Flexibility and Authority: Exploring Negotiation as an Interaction Strategy for Healthy Sleep Behaviors. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. 401–415.

[64] Konstantin R. Strömel, Stanislas Henry, Tim Johansson, Jasmin Niess, and Paweł W. Woźniak. 2024. Narrating Fitness: Leveraging Large Language Models for Reflective Fitness Tracker Data Interpretation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 646, 16 pages. https://doi.org/10.1145/3613904.3642032

[65] Keisuke Takano and Yoshihiko Tanno. 2009. Self-rumination, self-reflection, and depression: Self-rumination counteracts the adaptive effect of self-reflection. *Behaviour research and therapy* 47, 3 (2009), 260–264.

[66] Rob Tiffany. 2003. Data Manipulation Language. In *SQL Server CE Database Development with the. NET Compact Framework*. Springer, 139–179.

[67] Dominicus Tornqvist. 2014. exploratory Play in simulation sandbox Games: a review of What We Know about Why Players act crazy. *International Journal of Game-Based Learning (IJGBL)* 4, 2 (2014), 78–95.

[68] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *science* 211, 4481 (1981), 453–458.

[69] Ruth Van Veelen, Sabine Otten, Mara Cadinu, and Nina Hansen. 2016. An integrative model of social identification: Self-stereotyping and self-anchoring as two cognitive pathways. *Personality and social psychology review* 20, 1 (2016), 3–26.

[70] Garth Wells and Garth Wells. 2001. DML: Inserting, Updating, and Deleting Data. *Code Centric: T-SQL Programming with Stored Procedures and Triggers* (2001), 133–148.

[71] Sang Michael Xie, Aditi Raghunathan, Percy Liang, and Tengyu Ma. 2021. An explanation of in-context learning as implicit bayesian inference. *arXiv preprint arXiv:2111.02080* (2021).

[72] Rayoung Yang, Eunice Shin, Mark W. Newman, and Mark S. Ackerman. 2015. When fitness trackers don't 'fit': end-user difficulties in the assessment of personal tracking device accuracy. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 623–634. https://doi.org/10.1145/2750858.2804269

[73] Xinyue Zhou and Ding-Guo Gao. 2008. Social support and money as pain management mechanisms. *Psychological Inquiry* 19, 3-4 (2008), 127–144.

A APPENDIX A

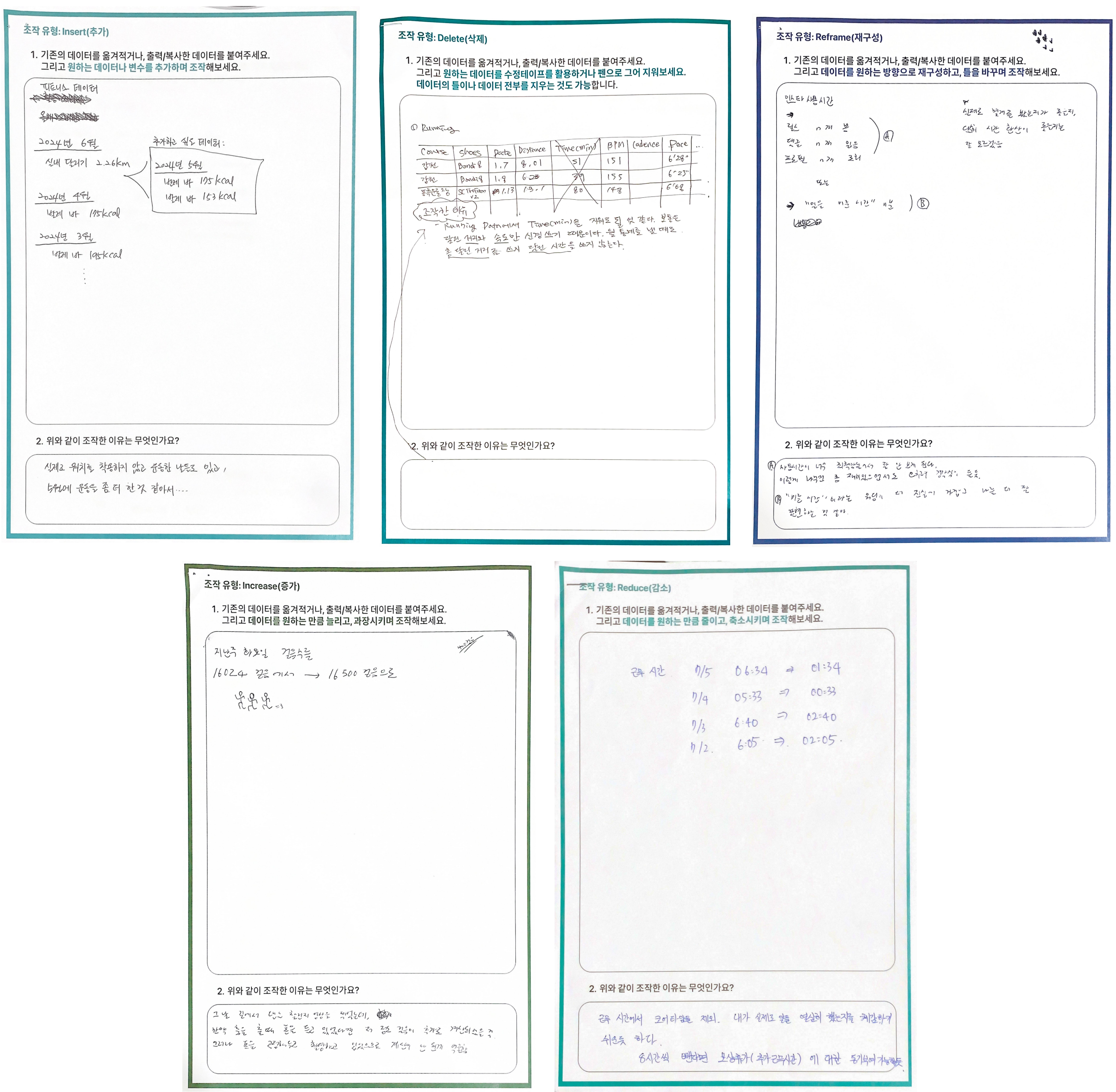
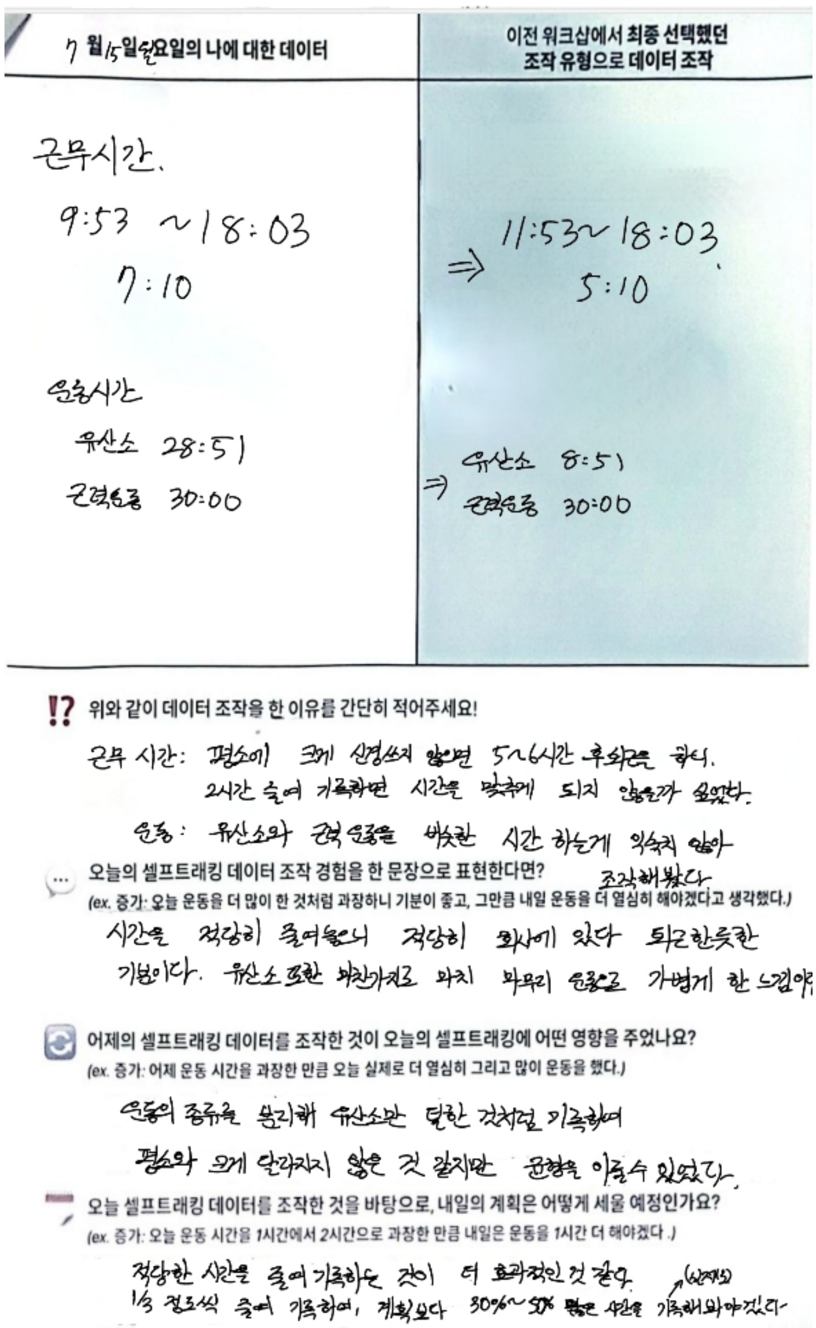


Figure 5: Original versions of paper worksheets in the exploratory workshops. The translated version is located in Figure 3 of the main text. We would like to note that the exploratory workshops were conducted using the term "Reframe", but we later updated to "Transform" to better align with the core idea we wanted to convey. While "Reframe" focuses on changing perspectives or interpreting data differently without modifying the original data, "Transform" more clearly reflects the concept of directly changing the structure or value of the data. Although the term has changed, the core idea and purpose remain the same.

B APPENDIX B



7 월 15 일 월요일의 나에 대한 데이터	이전 워크샵에서 최종 선택했던 조작 유형으로 데이터 조작
근무시간. 9:53 ~ 18:03 7:10	⇒ 11:53 ~ 18:03 5:10
운동시간 유산소 28:51 근력운동 30:00	⇒ 유산소 8:51 근력운동 30:00

!? 위와 같이 데이터 조작을 한 이유를 간단히 적어주세요!

근무 시간: 평소에 크게 신경쓰지 않으면 5~6시간 후 퇴근을 한다. 2시간 줄여 기록하면 시간을 맞추게 되지 않을까 싶었다.

운동: 유산소와 근력 운동을 비슷한 시간 하는게 익숙치 않아 조작해봤다.

오늘의 셀프트래킹 데이터 조작 경험을 한 문장으로 표현한다면?

(ex. 증가: 오늘 운동을 더 많이 한 것처럼 과장하니 기분이 좋고, 그만큼 내일 운동을 더 열심히 해야겠다고 생각했다.)

시간을 적당히 줄여놓으니 적당히 회사에 있다 퇴근한듯한 기분이다. 유산소 또한 줄여가지고 마치 마무리 운동으로 가볍게 한 느낌이다

어제의 셀프트래킹 데이터를 조작한 것이 오늘의 셀프트래킹에 어떤 영향을 주었나요?

(ex. 증가: 어제 운동 시간을 과장한 만큼 오늘 실제로 더 열심히 그리고 많이 운동을 했다.)

운동의 종류를 분리해 유산소만 덜한 것처럼 기록하여 평소와 크게 달라지지 않은 것 같지만 균형을 이룰수 있었다.

오늘 셀프트래킹 데이터를 조작한 것을 바탕으로, 내일의 계획은 어떻게 세울 예정인가요?

(ex. 증가: 오늘 운동 시간을 1시간에서 2시간으로 과장한 만큼 내일은 운동을 1시간 더 해야겠다.)

적당한 시간을 줄여 기록하는 것이 더 효과적인 것 같다. 1/3 정도씩 줄여 기록하여, 계획보다 30%~50% 많은 (실제로) 시간을 기록해봐야겠다.

Figure 6: Original version of paper worksheet in the one-week field trial. The translated version is located in Figure 4 of the main text.