Understanding Barriers of Missing Data in Personal Informatics Systems

Nannan Wen[1(✉)], Aditi Mallavarapu[2], Jacob Biehl[1], Erin Walker[3], and Dmitriy Babichenko[4]

[1] Department of Computer Science, School of Computing and Information, University of Pittsburgh, Pittsburgh, USA
{naw66,biehl}@pitt.edu

[2] Digital Promise Global and University of Pittsburgh, Pittsburgh, USA
amallavarapu@digitalpromise.org

[3] School of Computing and Information and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, USA
eawalker@pitt.edu

[4] Department of Informatics and Networked Systems, School of Computing and Information, University of Pittsburgh, Pittsburgh, USA
dmb72@pitt.edu

Abstract. Advanced personal informatics (PI) tools enable users to collect and reflect on a wide range of personal data. Researchers consider missing data, or discontinuous (sparse) data caused by device malfunctions or human errors, an important barrier for adopting PI tools in their daily routines. While a lot is known about why missing data occurs, less is known about its impact on user reflection or how tools can be designed to mediate/reduce its negative effects in PI systems. In this work, we focused on exploring the importance and impact missing data has on user reflection and extracting insights to improve the design of PI reflection tools. We present a semi-structured interview to investigate the impact of missing data on users' daily usage on two user groups, trainees and maintainers. We then provide design implications for incorporating visualization of estimated data (synthetic data) in the reflection stage, as a potential solution to the missing data problem. In this work, we provided data-driven implications for the design of future PI tools to help users reflect upon and mitigate missing data in their tracking activities.

Keywords: Personal Informatics (PI) · missing data · visualization · synthetic data

1 Introduction

Personal Informatics (PI), also referred to as "living by numbers", "quantified self", "self-surveillance", "self-tracking", and "personal analytics", is defined as "those [tools] that help people collect personally relevant information for the

A. Tsanas and A. Triantafyllidis (Eds.): PH 2022, LNICST 488, pp. 603–618, 2023.
https://doi.org/10.1007/978-3-031-34586-9_40

purpose of reflection and gaining self-knowledge" [20]. PI data have been used for tracking health through diet [1,35,36], exercise [23,24], changes in moods [2], sleep [25], menstrual cycles [14], and other physical and psychological activities. Extensive prior research has examined the use and utility of these technologies to improve behavior [6,21,27], understand differences in engagement and reflection on personal data [27], and build descriptive and predictive models of specific events and activities [13,20]. Research has also sought to expand what can be quantified, exploring new use domains [4,8,11,12,29]. In this work, we broadly refer to PI tools as the set of technologies and approaches that include wearable devices, systems, and applications that participants use to collect data on their personal behaviors and activities.

While there is great promise for PI-driven interventions, many practical obstacles impact users' ability to *collect* and *reflect* upon PI tools and associated data. For example, Li et al. [20] note three main barriers to data *collection*: 1) tools (devices) are not around when symptoms happen, 2) users forget to record activities, and 3) devices and applications lack necessary accuracy for helpful measurement. In *reflection*, the main obstacles include the sparsity of data and the ability to interpret data and resulting summaries, visualizations, and recommendations.

These obstacles have been shown to have a significant impact on adherence, interest, and trust in PI [10,13]. For example, Choe et.al. [6], who studied the practices of the Quantified Self movement, reported that participants valued tools (devices) that could capture comprehensive, granular information about their activities and expressed frustration when collected data was not accurate. Rapp et.al. [27] found that missing data might cause the users to mistake what activities were captured. Thomas et.al. [16], in their study of long-term use of wearable devices, describe that participants spent extra effort to ensure that they had their tools (devices) with them before their workout. They worried they would not get proper credit for their activities without the devices. Identifying incomplete data can also impact users' affect; Epstein et al. noted that participants felt guilty when a menstrual tracking application's interface noted missing data [12].

The literature, overall, notes that missing data plays a significant role in defining the value of PI [6,10,12,16,19,27], but few of them focused on resolving the effects of missing data. In this work, we conduct interviews in the hope of better understanding missing data's impact on the reflection stage and propose design implications to mediate its negative effect.

2 Related Work: Systems that Support Refection in PI

The most widely used definition of reflection in PI follows Schön's [31] reflection-in-action and reflection-on-action [3]. Systems that support reflection-in-action provide feedback during the activity. This concept is widely adopted in applications (e.g., [7,26,33]) that aim to promote physical health and encourage a healthy lifestyle using different methodologies. For example, commercial products like Fitbit and iWatch give real-time feedback when walking, running, or cycling. The Ally+ app [26] is an academic prototype that acts as a chat-based digital coach to deliver in-the-moment interventions to motivate participants to achieve their step goals. Systems supporting the concept of reflection-on-action usually provide feedback after the activity ends. Fish'n'Steps [22] is an example of such a system that promotes reflection-on-action; it is a game that links a player's daily step count to the growth of a virtual character, a fish, and a tank, to encourage physical activities. In the same line of work, the UbiFit [9] is another educational system designed to improve physical activity by using positive reinforcement based on past behaviors. In addition, Visualized Self [5] is a web-based system that supports deeper-level self-reflection through multiple data streams and visual data exploration using participants' historical data. Finally, Habito [17] is an android application that utilizes textual feedback on participants' activities to study how users engage with activity trackers.

While there are many different techniques to facilitate reflection, a specialized field of "personal visualizations" aims to present personally relevant information that promotes actionable insights and subsequent changes in behavior. Among those visualizations designed to increase awareness and encourage behavior change of self-trackers(e.g., [5,17,22]), many of them provided visualizations in the form of dashboards and supported simple interactions to explore the data, employing timeline metaphors to present events chronologically. For instance, Lifestreams [18] is an analytical tool to extract specific behavioral indicators and inferences from linear, interactive visualizations. Likewise, Moushhumi et al. [34] visualize time-series data to design just-in-time adaptive stress management interventions. These tools provided a valuable overview of trackers' activities.

Even though a considerable number of systems are available to aid data reflection, missing data was not considered a principle design element of their tools to the best of our knowledge. Nevertheless, it is well-recognized that missing data is an important barrier to PI. Two models have been proposed in the past two decades to study user behaviors through Personal Informatics systems. In the stage-based model, Li et.al. [20] divide the process of personal informatics into an iterative stage of preparation, collection, integration, reflection, and action. Epstein et.al. [13] grew their model by incorporating the perspective of lived informatics [28]. The interrelationship between stages in the lived informatics model is more complex than in the stage-based model due to the iterative nature between stages and substages. The lived-informatics model [28] includes the process of deciding, selecting, tracking & acting, and lapsing. Tracking & acting were further divided into an iteration of the collection, reflection, and integration process due to frequently switching and abandoning tools and periodically reviewing or reflecting on their data.

In Li's model, due to the cascaded dependency between the different stages, starting from the collection stage, missing and inaccurate data prevent people from going to the next stage or continuing the tracking activity. In the reflection stage, in both of this models [20,21], the discontinuous and missing data further prevent users from transitioning to the next stage, the action stage, through the lack of rich insight. In the lived-informatics model, Epstein et.al. [13] also identified that the reasons for stopping to track and starting again are the same barriers that Li et.al. [20] pinpointed. Even though researchers suggest tools should collect as much data as possible [20] to mediate the effect missing data has in the reflection stage, missing data are still present and causing problems.

3 Method

3.1 Research Design

Our study sought to understand the role of missing data and its subsequent impact on PI usage, utility, and related behaviors. Specifically, we sought to answer the following research questions: What are users' expectations of consistency and completeness in PI data? Does missing PI data conflict with these expectations? How can the design of PI tools be extended to help users reflect upon and mitigate missing data in their tracking activities?

To answer our research questions, we conducted semi-structured interviews. The participants need to own or have consistent access to PI systems, such as Fitbit, iWatch, Garmin, Strava, etc., for at least three months and have reviewed their systems' dashboard(s) and report(s) for self-reflection at least once a month. Twenty individuals participated in the study. In addition, a brief survey collected demographic information and had users rate (Likert scale) the impact of missing data on their PI goals and overall attitude towards missing PI data.

3.2 Participants

We sent our recruitment email to private training groups, Panther Cycling Club and Club Triathlon of Pittsburgh. We also sent out the recruitment questionnaire to a crowdsourcing website, Prolific, to recruit participants who have been using PI tools. As a result, 314 people responded to the questionnaire. We applied the inclusion criteria, which are 1) used PI tools for more than three months, 2) reflected on their data at least once a month, 3) collected data towards a specific goal, and randomly selected participants across the three recruitment sources. As a result, 20 people (11 female, 9 male) participated in the study, and the participants were aged 19 - 56 years, M = 32.875, SD = 9.34, and resided in the US. The number of participants was informed by the Qualitative Research and Saturation Criteria [30]; we concluded the surveys when no new information emerged from later interviews. As represented in Table 1, all participants had relatively extensive tracking experience.

3.3 Procedure

The design of this study was approved by our University's institutional review board (IRB). The potential participants who took the pre-screening Qualtrics survey received an email to schedule a virtual interview for up to an hour. This work was conducted during the COVID-19 pandemic; interviews were conducted over Zoom. We recorded the audio of the interviews. In the interview, we first gained informed consent and then engaged in a series of questions to have the participants describe the PI tools they used, their motivation, and how they used them. We then asked them to show us how they reflected on their data and how missing data would occur and influence this process. Then, we asked them to express their attitudes toward missing data in descriptive sentences and on a 5-point Likert scale. Finally, we asked them for suggested methods to mediate missing data. We compensated the participants $15 for their participation.

3.4 Data Preparation and Analysis

The audio files from the interview were imported and machine transcribed using the cloud-based platform Atlas.ti. Two researchers (both authors on the paper) used an open coding technique to identify themes and trends in the responses, and proceeded to identify relationships among the codes (axial coding [32]). Then the first researcher and a third researcher (not an author on the paper) separately read and coded 10% of the transcripts using the themes identified previously. Their initial inter-rater reliability (percentage agreement) was 0.769; which is higher than the expected agreement by mere chance, 0.56, proposed by Krippendorff ([15], p. 224–226). The first and the third researcher discussed discrepancies and updated the existing codebook. Afterwards, the first and a fourth researcher (not an author on the paper) coded the rest of the transcripts, and the inter-rater reliability by Krippendorff's alpha was 0.73, and the agreement was 0.99.

4 Results

Our study found notable user behaviors and perceptions in situations where data are missing in personal informatics tools. Specifically, several themes emerged that link missing data to capture and goal-tracking challenges. Within these themes, we observed two very distinct classes of users: ***trainees*** and ***maintainers***. Trainees are individuals with concrete, even professional athletic training goals who seek data to guide decisions to improve performance-related metrics (e.g., cardiac efficiency, time on interval). Maintainers have broad health improvement and maintenance goals. They seek data to track milestones (e.g., steps per day, active exercise hours). Given these very divergent goals, we present results from the perspective of *both* classes of users.

4.1 Usage of the PI Tools

To provide context, we first describe how trainees and maintainers used PI tools as part of their daily routines. A wide variety of PI tools were used across the 20 participants interviewed. Table 2 summarizes these tools, separating them between the two classes of users. We performed an analysis of each tool, assessing the presence of functionality along four dimensions: 1) data export, e.g., use of

Table 1. Participants demographics, tracking background, and frequency of reflection

PID	Age range	Gender	Occupation	Wearables	Duration	Category	Main motive of use	Frequency and types of reflection
P1	32–43	Male	Cycling coach	Garmin watch	13 years	Trainees	Training performance, analysis for cycling data, balance training load.	Reflect on physiological data after biking, deeper analysis on the weekend to check performance, modify training load, and analyze performance per season and year
P2	19–31	Female	Service	Apple watch	7 years	Maintainers	Maintain heart rate, maintaining weight, be active.	Check data after workouts, check meal data daily on the app for nutrition, reflect on weekends for meal summary
P3	32–43	Male	Teacher	Garmin Phoenix five	13 years	Trainees	Replicate best performance for the race, inform training schedule	Reflect on data after running, reflect on weekly data to analyze and identify best performance based on road conditions, pace, and duration
P4	32–43	Male	Software engineer	Kronos watch	1 year	Maintainers	Maintaining weight, improve health	Check to see if hit daily targets, check weekly data to see trends
P5	44–56	Female	Unemployee	VeryFit watch	2 years	Maintainers	Maintaining health, losing weight	Reflect on weekly data and daily data to see if hit targets
P6	19–31	Female	Desk job	Apple watch	> 6 years	Maintainers	Tracking activities, checking calories burned.	Multiple times a day to see if they hit daily targets, reflect three times a month to see summaries reflect, weekly to see if they hit weekly targets
P7	19–31	Male	Student	Suunto watch	3 years	Trainees	Tracking and analyzing cycling data, log exercises.	Multiple times a day, reflect after cycling for performance
P8	19–31	Female	Student	Garmin watch	2 years six mo	Maintainers	Tracking activities, hit exercise target.	Check during workouts, reflect at least once a day to see if hit targets
P9	32–43	Female	Unemployee	Fitbit	> 10 years	Maintainers	Tracking activities, log workouts.	Check data after workouts, check weekly to see exercise types and duration
P10	44–56	Male	Learning Consultant	Garmin watch	5 years	Trainees	Maintain certain pace during marathon	Reflect after running for pace, heart rate and road condition
P11	19–31	Female	Desk job	Garmin watch	5 years	Trainees	Training to advance calisthenics	Reflect three times a week for how many times reached the goal
P12	32–43	Female	Desk Job	Fitbit	5 years	Maintainers	Tracking exercises, maintaining weight, being active	Check multiple times a day during exercise, reflect once a day to see if hit targets
P13	19–31	Female	Unemployee	Fitbit	1 year	Maintainers	Tracking activities, be active	Checking data multiple times a day and during exercise, reflect daily to see if hit targets
P14	19–31	Female	Student	Apple Watch, Whoop band	4 years	Trainees	Training for Brazilian Jiu-Jitsu competition, improve performance	Reflect daily to check recovery score, inform the types of training the body is ready for the day, reflect weekly and monthly to check trends/patterns of the week&month
P15	19–31	Female	Student	Apple Watch	2.5 years	Maintainers	Tracking exercises, losing weight	Reflect daily to check if hit targets, check during exercise to see progress
P16	19–31	Female	Teacher	Apple Watch	2 years	Trainees	Training for weight lifting, building muscles	Checking after training for performance: duration, weight; reflect to inform training load
P17	32–43	Male	Pilot	Whoop band	1.5 years	Trainees	Training for hike Mount Rainier	Reflect after hiking for speed, heart rate, and altitude of the mountain
P18	44–56	Male	Maintenance technician	Garmin Vivoactive Three	8 years	Trainees	Training for running	Checking data for time and heart rate during training. Balance training load. Reflect after running for hill work, pace, distance, elevation, and duration
P19	32–43	Male	Teacher	Misfit Vapor X	4 years	Trainees	Training for cycling, making sure to hit endurance targets	Reflect weekly to inform training in the following week, reflect after cycling to check for performance
P20	19–31	Male	Wine tasting host	Apple Watch	> 1 year	Maintainers	Tracking workout, stay active	Checking data multiple times per day for progress, reflect weekly for summaries

data captured by the tool for use in another tool, 2) data visualization, e.g., visual depictions of activity and progress towards a defined goal, 3) data analysis, e.g., identify performance trends and calibrate future goals, and 4) social features, e.g., posting on social media a physical activity or accomplishment.

Table 2. PI wearables/applications used by participants. "●" means feature was being used by at least one participant, "○" means features was not being used by participants who reported using the wearables/applications, "—" means device does not have such a feature

Category	PI tools used	Usage of the PI tools			
		Data export	Data visualization	Data analysis	Social features
Trainees	Strava	●	●	●	●
	Pedometer	●	—	—	—
	Google Fit	●	○	○	—
	Power meter	●	—	—	—
	Whoop band	●	●	●	○
	Couch to 5k	●	—	—	—
	Suunto watch	●	○	●	○
	Apple watch	●	○	●	○
	Garmin watch	●	○	●	○
	Golden cheetah	—	●	●	—
	Bike computers	●	○	—	—
	GPS foot pedals	●	—	—	—
	Garmin Fenix watch	●	○	●	○
	Garmin Vivoactive watch	●	●	●	○
Maintainers	Lose it	○	●	○	—
	Virgin Pulse	○	●	○	—
	Apple watch	●	●	●	●
	Fitbit watch	●	●	○	●
	Veryfit watch	○	●	○	—
	Kronos watch	●	●	●	●
	My fitness pal	○	●	○	○
	My Maintainers Pal	○	●	○	○
	Apple health	○	●	○	●
	Garmin Connect	○	●	○	●
	Garmin Vivoactive watch	○	●	○	●

The analysis of feature presence included reflection on participants' explicit descriptions of use, our assessment of manufacturer marketing materials and technical manuals, and (if application-based) our independent installation and exploration. The four latter columns of Table 2 report the result of this analysis. A circle indicates we assessed the feature dimension to be present in the tool, and a dash, if assessed, is not present. A solid shading indicates at least one participant indicated using the feature. An empty circle indicated the feature was present, but observed no reported use among participants in the interviews.

We discovered several interesting patterns through this analysis. Most prominent, many of the *same* tools were used in different ways by the different classes of users. Specifically, ***trainees*** overwhelmingly favored the use of data export and data analysis functionalities over the built-in (or even *non-present*) data visualization features. The opposite use pattern was observed for ***maintainers***, who favored visualization features. We captured sentiment and motivation for these divergent behaviors in the interviews.

Interviews explored the lifecycle of personal informatics tools. As expected, tools are most commonly discontinued when users acquire new, more capable tools (e.g., upgrading to the latest fitness watch), resulting in missing data in their records.

Table 3. Reflective behaviors

Category	Motive of use	Data Usage (goal)	Frequency of use		Features used
			Collection	Reflection	
Trainees	Cycling Marathon Martial arts Coach others Weight lifting Cycling tournament Running tournament	Training power Training strength Training sprint Analyze performance Balance training load Inform training plan Optimize performance Improve performance Maximize abilities in running Replicate best performance	Upon training	Daily, upon training, weekly, monthly	Identify important moments; Training stress balance; Chronic training load; Elevation gain; Speed; Power; Heart rate zones; Cadence; Pace; Training intervals; Identify min/max value for each feature; Recovery score.
Maintainers	Educate self Log activities Improve health Maintain health Manage weight Maintain heart health	Map routes Plan meal Monitor Sleep Lose weight Maintain weight Check Mileage Regulate heart rate Check step counts Check Calories burned Monitor Sleep conditions Log Swimming hours	Daily	Weekly, monthly	Calories intake and nutrition; protein and carbohydrates in each meal; Active hours; Sleep duration per day; Total workout hours; Check completion rate; Check targets set for activities; Compare performance.

4.2 Users' Expectations of Consistency and Completeness in PI Data

Knowing how people utilize their PI tools for different motivations (see Sect. 4.1, Table 3 further discriminates participants' usage behaviors into ***trainees*** and ***maintainers***. We analyzed the features and frequency of use, during two stages: 1) collection, e.g., collecting sensory data using the PI tools, and 2) reflection, e.g., reflecting on the collected data to gain actionable insights for behavior change.

The analysis of the frequency and features used during reflection is extracted from participants' explicit descriptions of how they reflect. The last two columns report the result of this analysis. A notable distinction emerged between the different classes of users. ***Maintainers*** would regularly use and interact with the tools at the time of collection, often to confirm data collection and assess progress towards the goal. In contrast, the ***trainees*** were found less engaged, often checking the tools only once a day or periodically throughout the week to reflect on their activity trends and broader health maintenance goals.

Reflection behaviors with ***trainees*** often centered on *understanding the past* and *predicting the future*. For instance, P1, a trainee, stated the need to guide adjustments in a training routine, by reflecting on previous performance data: *"if somebody is losing races and their sprint is not as good as before, he/she will need more sprint work (recommended method of choice for cardiovascular exercise)"*. P3, also a trainee, noted he believes *"it helps you predict how fast you can actually race because you know exactly how long you can hold at a certain* *heart rate for"*. P13, a trainee again, stated the need to maintain a reasonable training stress balance: *"based on how recovered I am, this is how much strain I should put on my body"*.

In comparison, ***maintainers*** would often reflect to assess past behaviors against goals *knowing the present*. For example, P6, a maintainer, commented *"usually just checking how far I am from the target"*. In addition, 7 out of 10 maintainers *(P5, P6, P9, P4, P13, P8, and P20)* stated they mainly log how many exercises, or how many times for an exercise session, to see if they met the perfect week (hit targets every day), the perfect span, exercise, and weekly goals.

The value of precision and detail in PI also differed. ***Trainees*** were found to be interested in specific insights and would compare exact data across multiple periods or durations (e.g., daily, weekly, and monthly). For example, trainees use important moments, chronic training load, training intervals, and recovery scores to balance easy and heavy training. P3 noted that knowing the min/max of all these features helps him *"analyze from an analytical perspective about what my body is capable of"*, and claims one can even replicate a certain level of heart rate and pace to mimic past victories in a new tournament. On the other hand, ***maintainers*** focused on summative insights across the broad set of past activities. As the last column shows, features like active hours, sleep duration, total workout hours, etc., are oriented toward tracking trends and broad health goals.

4.3 Missing PI Data Conflict and Users' Expectations

The reflective behavior described in Sect. 4.2 is dependent on the PI data collected. We noted the presence of missing data in either of the stages to be problematic. For ***trainees***, activities that had no data prevented effective comparisons across activities or tracking specific performance metrics. For ***maintainers***, a lack of data could incorrectly imply long-term goals are not being achieved. Across both groups, the absence of key data was linked to the abandonment of the devices, indicating little tolerance for adapting to the functional limitations of the tools.

Both ***maintainers*** and ***trainees*** claimed they had experienced missing data when using the PI tools in our study. We analyzed the causal factors that lead to missing data and its effects during the lifecycle of personal informatics. Our analysis revealed two dimensions: 1) human reasons, e.g., missing data caused by participant error or behavior; 2) device reasons, e.g., missing data caused by malfunction or misconfiguration of wearables or applications. These dimensions apply in both groups.

Missing data caused by humans are: 1) *Forgetting to initiate data collection*, e.g., participants would forget to invoke the application or device to collect data for an activity. This could also include forgetting to annotate data or input manual entries. 2) *Forgetting to bring the device*, e.g., participants would forget to bring or wear the device for an activity. In some cases, like for P4, a ***maintainer***, and P8, a ***maintainer*** again, who forgot to bring their smartwatch on vacation,

Table 4. Quantifiable questions

Questions	Ratings
Q1. Please rate your frustration level when the data is missing	"Not at all"
Q2. Please rate your frustration level when the data is inaccurate	"Slightly"
Q3. Please rate your trust level towards your tracking data.	"Somewhat"
Q4. Please rate the level of influence missing data has on your goal.	"Moderately"
	"Extremely"

these gaps can be over many days. Most common, was forgetting to put the device back on after charging.

Missing data caused by devices include: 1) *Battery died*, a significant portion of the participants experienced battery issues during their activities. e.g., P8, ***maintainer***, claims *"sometimes I forget to charge my watch, it'll die in the middle of a workout or run"*. 2) *Malfunction or limitation of the device*, e.g., P6, a ***maintainer***, reported that her smartwatch did not count the swim laps. 3) *Syncing problem*, e.g., P7, a ***trainee***, stated that he uses multiple applications, and when syncing the data to other applications, data points were lost. 4) *Precision of the device*, all participants claimed that their device would often not capture activities at the correct level of precision, e.g., P4, a ***maintainer***, claimed that sometimes, when he was holding the wheel and driving, the smartwatch would pick up the vibration and count it as steps. P16, a ***trainee***, stated *"sometimes, if I'm walking around the class and talking really loud, my heart rate might spike a little bit, but it (smartwatch) may count that I'm actively working out, which I'm not"*.

The quantitative analysis on the 5-point Likert scale rating (questions shown in Table 4) showed that 20% maintainers and 40% trainees claimed missing data does not affect their goals (Q4, rating of "Not at all"), 60% maintainers and 40% trainees reported that missing data slightly influences their goals (Q4, rating of "Slightly"), 20% maintainers claimed missing data somewhat affects their goals (Q4, rating of "Somewhat"), and 20% maintainers reported missing data moderately or extremely influence their goals (Q4, rating of "Moderately").

20% maintainers and 30% trainees reported taking action when they noticed the missing data, and the same percentage of participants also claimed that if the data is missing, they will refer to their friends' data who were on the same route to estimate their PI data. Furthermore, 20% of the trainees would estimate the data based on how their body feels. In the reflection stage, if participants noticed the tools did not capture their hours/activities during the day, they would redo the workout to make up for the missed data. For those estimated data, they reflect on it with no difference from regular data, but some participants also claimed that they prefer no data to inaccurate data under some circumstances, e.g., P18, a ***trainee***, claimed *"The estimated calories burned are accurate to about 90%, which is fine. But for the estimated heart rate? No, absolutely not. I would rather the app (Google Fit) tell me it did not capture it (heart rate)."*

Interestingly, having data towards a goal would even impact some participants to repeat activities to properly capture the PI data. For instance, P8, a ***maintainer***, who claimed missing data somewhat affects her goal, commented that *"If I set 100 (minutes), and I had missing data and the watch shows I would have gotten 30 (minutes). I'll probably just run again with my watch, and then get more minutes, just so I can reach my goal (set) on the app"*.

Frustration with missing data (and, by consequence, inaccurate data) was also varied, 10% trainees and 20% maintainers claimed they felt somewhat frustrated when their data was missing (Q1, rating of "Somewhat"), 30% trainees and 20% maintainers stated their frustration towards missing data is moderate (Q1, rating of "Moderately"), and 30% trainees rated their frustration as extremely (Q1, rating of "Extremely"), an equal percentage of 20% in both groups expressed as slightly (Q1, rating of "Slightly"), and 10% trainees and 30% maintainers said they were not at all frustrated about missing data (Q1, rating of "Not at all"). All participants acknowledged that the data collected had some discrepancies, and it is not technically practical to have data with 100% accuracy. P5, a ***maintainers***, commented *"I like things to be accurate, I won't say I'm a number person per se, but I like things organized and proficient, so it (inaccurate data) did bother me at first"*. While acknowledging PI data is not 100% accurate, overall levels of trust in collected PI data were high. 50% trainees and 30% maintainers trusted their data extremely (Q3, rating of "Extremely"), 30% trainees and 50% maintainers trusted moderately (Q3, rating of "Moderately"), and 20% trainees and maintainers in each group trusted the data at a somewhat level (Q3, rating of "Somewhat").

From this analysis, we noticed that although missing data impact participants negatively in both groups, there are minor differences regarding the level of effects. For example, participants in the trainees' group tend to be more frustrated when their data is missing. On the other hand, the maintainers' group is more frustrated if the data is inaccurate.

5 Implications

From the semi-structured interview, we found different usage patterns for trainees and maintainers; we identified that missing data causes conflicts. We also elaborated on how their expectations were different between these two groups.

Our results contextualized the impact forgetting to wear or charge devices has on the value and utility of PI tools. This is a *fundamental* design limitation of PI technology – there is no clear technology horizon where these tools and devices automatically charge and attach to users. Designing *for* these events is critical to the long-term use and the utility of PI. Thus, in this section, we address the question: "how can the design of PI tools be extended to help users reflect upon and mitigate missing data in their tracking activities?"

Here we propose to take synthetic data as a principle when designing PI tools. By synthetic data, we meant to provide a visual representation in helping users to distinguish missing data vs. no data collected. And assist users in estimating their missing data during the reflection process.

Our results pointed to several key implications on how PI tools could be improved to embrace synthetic data in the user experience. Including synthetic data (missing data estimation) in the design principle of PI means the tool should not only consider where to include it in the tool, how to implement it based on different motivations, but also on how to represent it in the tool.

5.1 Usage Behavior

Our study showed three usage behaviors: 1) *understanding the past*, 2) *knowing the present*, and 3) *predicting the future* (see Sect. 4.2). Usage behaviors towards collected data are essential to self-reflection and finding actionable insights for behavior change. Our analysis indicated that current tools lack support for integrating that three usage behaviors. As a result, participants use separate PI tools for different use cases. This implies tools that enable synthetic data should consistently represent: 1) where synthetic data exists of weekly, monthly, and yearly data; 2) how synthetic data is used to estimate goal tracking; 3) control the use of synthetic data to discover patterns used to inform future behavior towards specific events or goals. All three usage behaviors are interconnected; a consistent representation of how missing data is represented and interpreted can reduce fallacious insight and increase users' confidence in their data.

5.2 Sensitivity to Missing Data

The result showed diverse levels of tolerance to missing data by participants (see Sect. 4.3). For example, some participants claimed they could accept one or two days of missing data per week. Some tolerate only a few hours per day. Participants consistently expressed tolerance for missing data within the context of interference with tracking goals, trends, and specific activities. This finding suggests no easy solution to meet missing data sensitivity through a single mitigation approach. Feedback from participants suggests three design approaches: 1) Allow the user to fill in the missing data manually; ideal for small gaps where a specific activity was performed (e.g., an outdoor run). 2) Use algorithmic mediation to estimate gaps in data based on prior data collection; ideal for large data gaps in daily, repeated activities (e.g., determining steps taken on a routine evening walk). 3) Use algorithmic mediation and guiding user interfaces to help users refine estimates based on known conditions of the missing data; ideal for larger gaps where a specific activity or set of activities was performed) (e.g., determine the performance of a long-distance run based on knowing the start and stop times and locations). Future work is needed to understand the effectiveness of each approach, and under what conditions each approach is preferred by users.

5.3 Visual Representation of Missing Data

For all the PI tools reported in use by participants in our study (See Table 2), the analysis indicated the tools handle missing data as the absence of data.

Exercise or activity could have been performed, but because data on the event was not collected, the tools often convey to the user *no activity* was conducted. This is misleading and can incorrectly impact a user's personal health goals. Tools should provide a better visual distinction between when data is missing and when activity is simply low. Identifying gaps when a device was worn can help users contextualize the data in the view, allowing them to reflect better on their physical activities. Further, visualizing gaps can be a natural opportunity in the user experience to allow users to estimate missing data manually or to enable an algorithm to generate heuristically derived synthetic data.

5.4 Trust in Data

Our analysis showed that participants largely believe the PI data collected is accurate (see Sect. 4.3). Trust towards the PI tools is critical to adherence and continued use. Participants would identify other sources for direct or interpretive comparison to further verify this trust. For instance, participants would compare data of individuals that also performed the same activity, examine the map while they ran or walked on a specific route in real-time, and estimate data reflecting on the tiredness of their body. These behaviors imply tools that enable synthetic data creation should enable users to: 1) observe synthetic data within the context of plausible and prior representative behaviors to assure synthetic data is grounded in primary data (e.g., comparison to previous five similar activities); 2) group and label synthetic data as specific, personally relevant events (e.g., labeling 'run with John'); and 3) be able to manipulate and revise algorithmic estimates based on intuition or preference in how missing data should be created.

6 Conclusion

This work considered a new perspective (missing data) on personal informatics and investigated how missing data impacts user behaviors and perceptions in authentic, everyday use. The work extended the community's understanding of the circumstances that cause missing data and how those circumstances influence collection and reflection behaviors. Our analysis demonstrated the issues caused by missing data at the level of individuals and technical issues. In addition, it contributed new insights into how personal informatics tools should adapt to support synthetic data and provided insights to guide how they can do it.

During the analysis, we identified two distinct groups, trainees and maintainers. We found trainees mostly use their PI tools to *understand their past* and *predict the future*. In contrast, the maintainers often use their PI tools to *know the present*. For both groups, we have provided a comparative result on missing data's influence on the lifecycle of personal informatics, and the implications capture lessons from the perspective of both groups. The resulting analysis point to key limitations of current tools, which is the lack of representation for missing data, and outlined design guidelines for future tools to improve the user experience with PI data. For future work, we would need to implement different methodologies based on the implementations to generate and represent synthetic data, conduct user studies to investigate the effectiveness of each approach, and determine under what conditions each approach is preferred.

References

1. Agapito, G., et al.: DIETOS: a recommender system for adaptive diet monitoring and personalized food suggestion. In: 2016 IEEE 12th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 1–8. IEEE (2016)
2. Alvarez-Lozano, J., et al.: Tell me your apps and i will tell you your mood: correlation of apps usage with bipolar disorder state. In: Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments, pp. 1–7 (2014)
3. Baumer, E.P., Khovanskaya, V., Matthews, M., Reynolds, L., Sosik, V.S., Gay, G.: Reviewing reflection: on the use of reflection in interactive system design. In: Proceedings of the 2014 Conference on Designing Interactive Systems, pp. 93–102 (2014)
4. Choe, E.K., Lee, B., Kay, M., Pratt, W., Kientz, J.A.: SleepTight: low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 121–132 (2015)
5. Choe, E.K., Lee, B., Zhu, H., Riche, N.H., Baur, D.: 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, pp. 173–182 (2017)
6. Choe, E.K., Lee, N.B., Lee, B., Pratt, W., Kientz, J.A.: Understanding quantified-selfers' practices in collecting and exploring personal data. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1143–1152 (2014)
7. Choi, W., Park, S., Kim, D., Lim, Y.K., Lee, U.: Multi-stage receptivity model for mobile just-in-time health intervention. Proc. ACM Interact. Mobile, Wearable Ubiquit. Technol. **3**(2), 1–26 (2019)
8. Chung, C.F., et al.: Boundary negotiating artifacts in personal informatics: patient-provider collaboration with patient-generated data. In: Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, pp. 770–786 (2016)
9. Consolvo, S., et al.: Flowers or a robot army? Encouraging awareness & activity with personal, mobile displays. In: Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 54–63 (2008)
10. Consolvo, S., McDonald, D.W., Landay, J.A.: Theory-driven design strategies for technologies that support behavior change in everyday life. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 405–414 (2009)
11. Epstein, D.A., Jacobson, B.H., Bales, E., McDonald, D.W., Munson, S.A.: From " nobody cares" to "way to go!" a design framework for social sharing in personal informatics. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 1622–1636 (2015)
12. Epstein, D.A., et al.: Examining menstrual tracking to inform the design of personal informatics tools. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 6876–6888 (2017)

13. Epstein, D.A., Ping, A., Fogarty, J., Munson, S.A.: A lived informatics model of personal informatics. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 731–742 (2015)
14. Fox, S., Howell, N., Wong, R., Spektor, F.: Vivewell: speculating near-future menstrual tracking through current data practices. In: Proceedings of the 2019 on Designing Interactive Systems Conference, pp. 541–552 (2019)
15. Friese, S.: ATLAS. ti 8 windows-inter-coder agreement analysis copyright 2019 by atlas. ti scientific software development GmBH, berlin. all rights reserved. manual version: 652.20190425. Updated for program version: 8.4 (2019)
16. Fritz, T., Huang, E.M., Murphy, G.C., Zimmermann, T.: Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 487–496 (2014)
17. Gouveia, R., Karapanos, E., Hassenzahl, M.: 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, pp. 1305–1316 (2015)
18. Hsieh, C.K., et al.: 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, pp. 1–13 (2013)
19. Lazar, A., Koehler, C., Tanenbaum, T.J., Nguyen, D.H.: Why we use and abandon smart devices. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 635–646 (2015)
20. Li, I., Dey, A., Forlizzi, J.: A stage-based model of personal informatics systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 557–566 (2010)
21. Li, I., Dey, A.K., Forlizzi, J.: Understanding my data, myself: supporting self-reflection with ubicomp technologies. In: Proceedings of the 13th International Conference on Ubiquitous Computing, pp. 405–414 (2011)
22. Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H.B.: Fish'n'Steps: encouraging physical activity with an interactive computer game. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 261–278. Springer, Heidelberg (2006). https://doi.org/10.1007/11853565_16
23. Martin-Niedecken, A.L., Rogers, K., Turmo Vidal, L., Mekler, E.D., Márquez Segura, E.: ExerCube vs. personal trainer: evaluating a holistic, immersive, and adaptive fitness game setup. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–15 (2019)
24. Miller, A.D., Mynatt, E.D.: StepStream: a school-based pervasive social fitness system for everyday adolescent health. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2823–2832 (2014)
25. Min, J.K., Doryab, A., Wiese, J., Amini, S., Zimmerman, J., Hong, J.I.: Toss'n'turn: smartphone as sleep and sleep quality detector. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 477–486 (2014)
26. Mishra, V., Künzler, F., Kramer, J.N., Fleisch, E., Kowatsch, T., Kotz, D.: Detecting receptivity for mHealth interventions in the natural environment. Proc. ACM Interact. Mobile, Wearable Ubiquit. Technol. **5**(2), 1–24 (2021)
27. Rapp, A., Cena, F.: Personal informatics for everyday life: how users without prior self-tracking experience engage with personal data. Int. J. Hum Comput Stud. **94**, 1–17 (2016)

28. Rooksby, J., Rost, M., Morrison, A., Chalmers, M.: Personal tracking as lived informatics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1163–1172 (2014)
29. Saksono, H., Castaneda-Sceppa, C., Hoffman, J.A., Seif El-Nasr, M., Parker, A.: StoryMap: using social modeling and self-modeling to support physical activity among families of low-SES backgrounds. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2021)
30. Saunders, B., et al.: Saturation in qualitative research: exploring its conceptualization and operationalization. Qual. Quant. **52**(4), 1893–1907 (2018)
31. Schon, D.A.: The Reflective Practitioner: How Professionals Think in Action. Basic books, New York (1983)
32. Scott, C., Medaugh, M.: Axial coding. Int. Encycl. Commun. Res. Meth. **10**, 9781118901731 (2017)
33. Sellak, H., Grobler, M.: MHealth4U: designing for health and wellbeing self-management. In: 2020 35th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW), pp. 41–46. IEEE (2020)
34. Sharmin, M., et al.: 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, pp. 505–516 (2015)
35. Theodoridis, T., Solachidis, V., Dimitropoulos, K., Gymnopoulos, L., Daras, P.: A survey on AI nutrition recommender systems. In: Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments, pp. 540–546 (2019)
36. Franco, R.Z.: Online recommender system for personalized nutrition advice. In: Proceedings of the Eleventh ACM Conference on Recommender Systems, pp. 411–415 (2017)