Reflection in Theory and Reflection in Practice: An Exploration of the Gaps in Reflection Support among Personal Informatics Apps

Janghee Cho
Department of Information Science
University of Colorado Boulder
Boulder, Colorado, USA
janghee.cho@colorado.edu

Tian Xu
Department of Information Science
University of Colorado Boulder
Boulder, Colorado, USA
tixu6250@colorado.edu

Abigail Zimmermann-Niefield
Department of Computer Science
University of Colorado Boulder
Boulder, Colorado, USA
abigail.zimmermannniefield@colorado.edu

Stephen Voida
Department of Information Science
University of Colorado Boulder
Boulder, Colorado, USA
svoida@Colorado.edu

ABSTRACT

Personal informatics (PI) systems have been developed to support reflection. While reflection is considered an indispensable activity in PI use, how and when reflection occurs is still under-studied. In this paper, we present an analysis of the interactive features of 123 commercial PI apps, revealing that reflective practices are unevenly supported. The lack of features that encourage user-driven reflection, scaffolding for setting goals and configuring data collection and presentation, and consideration of wider implications stand to limit meaning-making and frustrate nuanced insight generation based on lived experiences. Based on our findings, we discuss how reflection is currently misrepresented in personal informatics tools, identify and characterize the gaps between theoretical research on reflection and interface features in current apps, and offer suggestions about how reflection could be better supported.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

personal informatics, reflection, self-tracking, agency, mobile apps

ACM Reference Format:
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 *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 23 pages. https://doi.org/10.1145/3491102.3501991

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04.
https://doi.org/10.1145/3491102.3501991

1 INTRODUCTION

Personal informatics (PI) refers to a class of systems that help people collect data about and reflect on their experiences or behaviors to gain self-knowledge and induce positive behavior change(s) [96]. PI tools (e.g., wearable trackers, mobile or web tracking apps) have been increasingly adopted and used for recording large volumes of personal data. They have become a popular way to track physical behaviors or emotional states in the course of individuals' everyday lives, and they provide information that can support a wide range of personal goals [146].

PI research considers *reflection* an indispensable activity that enables individuals to generate insights for self-improvement and/or make lasting changes in behavior or mental schemas [16, 49, 100]. Although the notion of reflection is not a single-faceted concept, in a general sense, reflection is *a meaning-making process enacted through a conversation between a human being and their experiences, an object, or a situation.* [30, 42, 151]. Across many disciplines, reflection has long been acknowledged as a valuable and important practice to enrich lives; reflection allows people to engage in alternative ways of knowing in order to gain awareness and insights that empower them to make conscious decisions for both personal well-being and to advance societal values such as justice or environmental sustainability [152]. Thus, supporting reflection is viewed as an alternative way to overcome the dominant approach (i.e., persuasion) in behavior change interventions, helping users to maintain a sense of agency in making sense of their experiences rather than following instructions prescribed by others [13, 20].

Most PI research, including two broadly-cited PI models [53, 96] that aim to describe the role of reflection in the use of these tools, fall short in clearly articulating *how* and *when* reflective practices occur [14]. Reflection is usually taken for granted in the use of PI tools; as long as individuals are presented with data collected about their activity or behavior, PI systems have often been presumed to enable individuals to "gain insights" about their activity or behavior and to facilitate (potential) behavior change. Kersten-van Dijk et al. [84] called this assumption the *self-improvement hypothesis*, viewing a user as a self-motivated person who automatically reflects on their behavior by examining their data and subsequently changing their behavior based on what they've learned [119]. Without an explicit understanding of the definition and process of reflection, it is difficult to design technological interventions to scaffold effective reflective practices [14, 155]. Thus, a better understanding of how reflection can be stimulated to support a meaningful interaction with tracking tools is required.

While previous research provides conceptual frameworks of reflection as a resource for designing technologies [12, 17, 59, 155], little work has been done to characterize how these conceptual definitions can effectively be instantiated in system designs. Therefore, in this study, we propose a series of operational definitions of reflection by applying Fleck and Fitzpatrick's [59] conceptual definition of reflection to the domain of personal informatics, in order to ask the following research question: *how does the design of commercial PI apps support reflection?* To answer this question, we surveyed 123 commercially available PI apps, coding their interaction features as they relate to Fleck and Fitzpatrick's levels of reflection [59]. The ultimate objective of this research is to identify *blind spots* [152] (after) in existing implementations of personal informatics and expand the possible design space of these systems to more effectively bridge between a broader theoretical spectrum of reflection types and the practical issues involved in designing PI experiences.

Through this survey, we find that reflective practices in PI apps are unevenly supported. The lack of features that encourage user-driven reflection, scaffolding for setting goals and configuring data collection and presentation, and consideration of wider implications stand to limit meaning-making and frustrate nuanced insight generation based on lived experiences. Our research provides three main contributions. First, we extend Fleck and Fitzpatrick's theoretical framework of reflection to further operationalize each level of reflection in design features relevant to PI systems. Second, we discuss the ways in which reflective practices are—and are not—supported in current PI apps. Third, we provide implications about how reflection could be better supported in PI systems.

2 RELATED WORK

Personal informatics tools are primarily designed to capture longitudinal data and support reflection in generating insights for self-improvement and facilitating behavior change. Li et al. introduced a stage-based model of PI systems with an emphasis on the *collection* and *reflection* aspects of the systems that aid people in gaining self-knowledge [96]. Li and colleagues' model consists of a series of five stages— *Preparation*, *Collection*, *Integration*, *Reflection*, and *Action*—based on a goal-oriented perspective. Individuals move through each stage to obtain self-knowledge by looking at their data and subsequently changing their behavior based on what they learn from the interaction with the PI system [84]. Li et al. [96] note that each of these stages can be user-driven, system-driven, or a combination of the two, depending on whether a user or a system has a responsibility to perform each activity. They argued that designers should consider the tradeoffs between having the user and/or system be in control at each of these stages.

Complementing this influential model, Choe et al. [25] present three different ways for PI systems to collect data: (1) fully manual tracking, (2) fully automated tracking, and (3) semi-automated tracking. Choe and colleagues point out that reflective practices are also influenced by the type of data-collecting method(s) that are employed [28]. The fully-manual approach to tracking supports greater awareness of behavior because people are asked to input their data manually, and the action of entering data into the system forces users to engage with their tracked behavior. However, systems that employ fully manual tracking also incur a burden for users, leading to reduced convenience and increased abandonment [50]. To reduce users' burden, fully automated tracking may be considered as an alternative means for collecting data in PI systems. However, the automatic sensing approach reduces users' awareness of what data are collected and how they change over time [97]. As a compromise between these two approaches, Choe et al. [25] introduced the idea of *semi-automated* tracking, which they refer to as any combination of manual and automated tracking approaches. To successfully employ the semi-automated tracking mode in the PI tools, understanding what type of data is necessary to collect is essential [25]. Building on a semi-automated tracking approach, Kim et al. presented Omnitrack, a mobile PI tool that allows people to customize their tracking items depending on their tracking needs or preferences [88].

Niess and Wozniak introduced the Tracker Goal Evolution Model to describe how users set and track their goals [125]. To better understand long-term engagement with PI, they focused on users' goal practices to complement existing PI models. As Li et al.'s PI model ends in action, goal-setting theory [105] has been considered an effective strategy for encouraging behavior change (e.g., [33, 62, 67, 120]). Setting regular goals helps individuals dedicate themselves to achieving their goal(s) [104]. To keep individuals engaged with their goals, Gulotta et al. [73] argue that asking users to reflect on their goals (i.e., "periodic reflection") is important. They introduced two types of personalization to support the iterative refinement of goals: system-driven personalization (inferring users' goals based on collected data) and user-driven personalization (supporting the design of personal goals by the users themselves). For instance, machine learning methods are often applied to PI systems in order to generate system-driven recommendations that help users set their goals [62, 179]. Niess and Wozniak found that users reflect on their needs to translate their goals from qualitative to quantitative ones [125].

Rooksby et al. [146] extended the discussion about the use of PI systems by employing a "put people first then technology" approach to better understand the role of these technologies in users' everyday lives, resulting in a perspective that they refer to as *lived informatics*. Through the presentation of five styles of personal tracking (i.e., directive tracking, documentary tracking, diagnostic tracking, collecting rewards, and fetishised tracking), they argue that self-tracking technologies should support the diverse personal contexts and characteristics of lived experience rather than a single, goal-oriented trajectory. Failure to support diverse forms of use can result in an increase in abandonment [51] or introduction of ethical issues (e.g., marginalizing people from certain backgrounds [35, 157]). Mols et al. also argued that incorporating people's everyday life experiences into the design of technological intervention is crucial to support reflection [117]. To incorporate different modes in tracking, the need for customizability and flexibility in designing PI tools has been discussed (e.g., [5, 6, 38, 87, 107]). Although customizable features allow users to deeply engage with their experiences by supporting their agency in making sense of data, people may feel burdened by deciding how to track their data by themselves [5, 181].

Combining this lived informatics perspective with Li and colleagues' conventional PI model, Epstein et al. [53] proposed a new model for personal informatics. While the stage-based model focused primarily on the interaction between a PI system and its user, Epstein et al.'s lived informatics-informed model articulates a broader trajectory of self-tracking integration into everyday life by people with varied goals. Although the lived informatics model helps to provide a holistic view of the use of PI tools in everyday life, it is still unclear how people reflect on their experiences or behaviors when interacting with PI tools.

Reflection is often taken for granted in PI systems as long as users interact with the system as imagined by its designers. Given the fact that reflection is viewed as a naturally sequential process in PI system models (e.g., an inevitable consequence of interaction with the information visualization, the impetus for behavior change) [53, 96], individuals' reflective practices were understood as a mechanism of the information processing in PI tools, isolated from personal lived experiences. Although reflection could occur as long as output from PIs is displayed to users, this restrictive perspective of reflection (i.e., the information-processing metaphor [76]) may fail to capture all the ways reflection can be stimulated in the interaction with PI tools. Therefore, in this study, we focus on more closely understanding this reflection process, with the goal of better informing design to support diverse and situated reflection as part of self-tracking practice.

3 THEORETICAL BACKGROUND OF REFLECTION

Reflection is connected to multiple cognitive processes and can result in different outcomes, such as learning, behavior change, or gaining a critical perspective [59]. In the personal informatics model, for example, reflection is an impetus for behavior change [96], while in the reflective learning literature, reflection is an impetus for learning [18]. Although reflection is considered as a meaningful activity across different disciplines or practices, the notion of reflection remains ambiguous and contested [30, 59, 166]. This is because reflection is a multifaceted concept. Therefore, the definition or goal of reflection can be different depending on a theoretical framework in which genealogy or discipline is grounded. In this section, we develop the theoretical backdrop of our study, reviewing a range of foundational works on reflections with a primary focus on HCI.

Much of the HCI-oriented reflection literature traces the intellectual genealogies of reflection back to Dewey [42], who characterized reflection as an "active and deliberative cognitive process, involving sequences of interconnected ideas which take account of underlying beliefs and knowledge" (as paraphrased in Hatton and Smith [78]). The concept of *reflection-in-action* [151], introduced by Schön, has also recently been influential in HCI as a guide for developing systems to support reflection [14]. Schön [150] distinguishes between two types of reflection. *Reflection-on-action* is retrospective critical thinking—thinking back on what one has done in order to reconstruct the knowledge that was used [58]. *Reflection-in-action*, in contrast, is the embodiment of knowing-in-action—thinking about what one is doing while one is doing it, giving the reflective practitioner an opportunity to redesign what they are doing on the fly [151].

Fleck and Fitzpatrick present a taxonomy of five different levels of reflection as a resource for design based on the educational research literature [59]. The lowest level of reflection (R0) is referred to as a simple "description or statement about events *without* further elaboration or explanation" [59]. Although this level of reflection is considered as a part of the reflection stage in PI use [96], Fleck and Fitzpatrick's framework consider it to be mere revisitation of what one did, instead of constituting reflective practice. Their second level of reflection (R1) extends this idea of revisiting action, but includes prompts for individuals to articulate an explanation to justify those actions. Fleck and Fitzpatrick note that R1 is actually the first level of reflection to initiate reflective practice; therefore, we refer to R0 as *descriptive* reflection and R1 as *explanatory* reflection to distinguish between them. The third level of reflection (R2) is *dialogic* reflection, in which individuals examine the relationship(s) between two or more data points. The goal of dialogic reflection is to establish causality or correlation between one's previous experiences and his/her data. Next, *transformative* reflection (R3) helps individuals to develop a new perspective for reassessing their own orientation to perceiving, feeling, or acting [114]. If individuals become aware of alternative perspectives in level R2, then they can adapt those new perspectives as part of their mental schema or behavior in the R3 level. The last dimension is *critical* reflection (R4), which refers to reflection on aspects that transcend the immediate context (e.g., *social and ethical issues*). There seems to be a significant distinction between the lower levels of reflection (R0–R2) and higher levels of reflection (R3–R4) in this framework. While lower levels of reflection are situated in the displayed contexts (e.g., viewing data or visualizations of data), higher levels of reflection require an engagement with perspectives that transcend the immediate context.

To understand reflection as a multifaceted practice, we base our research on this framework because it best synthesizes conceptualizations of reflection with diverse literature on the practice of reflecting from an HCI perspective (e.g., [92, 100, 148]). Although both Dewey and Schön provide a solid theoretical basis of reflection, it is difficult for HCI researchers and practitioners to design features in computing systems by directly applying the breadth of these ideas about reflection. Moreover, a background in which these notions of reflection are discussed (i.e., John Dewey's reflection in a classroom setting, Donald Schön's reflective practices in a workplace) is quite different from technology use in everyday lives, even though their notion of reflection is still meaningful to discuss technology design in a conceptual manner. Bentvelzen et al. also noted that many theoretical frameworks of reflection were developed within the domains of specific professionals [17].

Other models of reflection employed in the HCI literature resonate with, if not duplicate the concepts contained in, Fleck & Fitzpatrick's work. Building on Mezirow's transformation theory [114], Mols et al. introduced the concept of *everyday life reflection* that is "all deliberate and critical thought processes concerning our day to day activities" [118, p.53]. Incorporating this perspective into the Fleck & Fitzpatrick model, one might expect that more time and multiple mechanisms may be required for supporting transformative reflection [155].

To critically reflect on not only oneself but also wider contexts, it is necessary to allow a user to form altruistic perspectives that transcend an ego-centric perspective. While Sengers et al. [152] used the word "critical" to define reflection, their notion of reflection is also aligned with both transformative reflection (R3) and critical reflection (R4) in the Fleck and Fitzpatrick framework. Sengers et al. viewed technology as a mediation that enables a designer to share their reflective concerns to make users reflect on not only their lives but also the impact of technology on their behavior(s).

4 METHODS

In this study, we conducted an iterative review of the reflection literature in HCI to construct a codebook linking levels of reflection with the design elements that commercial PI systems currently employ to instantiate and support different kinds of reflection. We then used this codebook, informed both by the literature and the interface elements we observed in commercially available mobile apps from Apple's App Store, to conduct a systematic review that assesses how deployed personal informatics apps support different levels of reflection based on our operational definitions.

Apps are considered as digital artifacts that are the product of "human decision-making, underpinned by tacit assumptions, norms and discourses already circulating in the social and cultural contexts in which app designers are generated" [108]. While app analysis is not a direct substitute for conducting human subjects research (e.g., experiments, surveys) to explore how end-users may or may not be appropriating PI systems' features for reflection in practice, doing so can still offer insights about dominant design approaches and to explore design spaces in an effort to improve the overall design and coverage of apps in a domain (e.g., [19, 37, 41, 139]). Our use of an app analysis method allowed us to see how apps prefigure a means for users to reflect on their own recorded data and to explore a new design space for reflection within PI apps.

4.1 Data Collection

We constructed our corpus spanning two data collection sessions. Initially, we compiled a corpus of mobile apps based on the list of "popular" apps on Apple's US App Store in June 2019[1]. We established our initial corpus from the most popular apps in the "Health & Fitness" category (240 total). It is only because many apps in this category incorporated functionality common to personal informatics. We then added 42 more PI-related non-"Health & Fitness" apps by searching for the keywords "track" and "tracking" from among the other categories of "most popular" apps in the App Store (see also Figure 1). These 42 additional entries included apps from the medical, finance, navigation, and book categories. Because our research experienced a pause during the COVID-19 pandemic, we refreshed our corpus to capture new and updated apps in the summer of 2021, verifying the inclusion of popular apps based not only on Apple's App Store rankings, but by additionally consulting a third-party app analytics list[2]. During this second data-collection pass, we added 53 more PI-related apps (across all categories on the App store) to the initial corpus.

Once we collected our initial corpus of tracking apps, we began to cull the list of apps to provide a tighter personal informatics focus based on the following heuristics. First, we excluded apps unrelated to self-tracking; for instance, some of the apps in our initial corpus were designed to track parcels, weather, or transportation schedules. We also excluded apps if they simply provided information (e.g., health education apps) without including any self-tracking features. Next, we excluded those apps that we would not be able to fully investigate: those that required payment, that did not offer a free trial period, that required registration in or association with a certain organization or company (e.g., apps tied to a particular health insurance provider or commercial weight-loss program), and that required purchase of an additional or auxiliary device, such as a smartwatch or external pedometer (e.g., Fitbit devices). This is because we aim to include apps having relatively low barriers to be selected and used by broader smartphone users. In other words, we attempted to include apps that might have a higher accessibility for potential PI users. Even though we excluded paid apps from our corpus due to the accessibility, we checked all paid plans from free apps (i.e., in-app purchase) to make sure if they provided additional features to support reflection. So when necessary, we created accounts and obtained free trials of "freemium" apps to ensure exploration of all reflection-oriented features. For the same reason, we also checked to see whether those apps remaining in our corpus were also available on Google's Play Store in order to ensure that our corpus did not embody a platform-specific bias, although we created our initial corpus from Apple's App Store; we removed those apps that were not available on both iOS and Android. Finally, we excluded lightly-reviewed apps—considering those with fewer than 1,000 posted reviews as outliers (the average number of reviews per app in our corpus is greater than 50,000). During our analysis phase, while we were collecting in-depth information about each of the apps, their functionality, and their level of reflection support, we discovered that eight apps had been removed from the App Store since we originally constructed our corpus; we excluded them from our reporting here, as well. Our final corpus has a total of 123 apps (see also Figure 1); a more detailed overview of this corpus appears in our supplementary materials.

The apps in this study reflect a diversity of tracking domains. However, three types (Physical Activity: 25%, Physical Activity & Food: 17.0%, Period: 12.1%) are the most dominant in our sample. As prior literature has shown, the most popular data that individuals reported tracking were physical activity (40%), food (31%), and weight (29%) [28].

4.2 Data Analysis

We iteratively constructed a codebook linking Fleck & Fitzpatrick's levels of reflection, the HCI literature on personal informatics, and those concrete design elements that we observed in the apps that we studied. We conducted qualitative coding of these apps' design features in two phases: initially, in the summer of 2019, and then with

[1]Retrieved from https://apps.apple.com/us/genre/ios-health-fitness/id6013, June 11, 2019. Our data scrape retrieved the app's title (name), short tagline, rating, category, and developer information.

[2]Retrieved from https://appfigures.com/top-apps/ios-app-store/united-states/iphone/top-overall, August 23, 2021

process and corpus development CHI2022.png

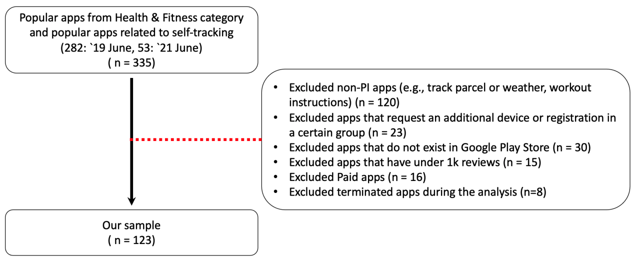


Figure 1: Sampling process and corpus development

Tracking domain	*N*	*% of corpus*
Physical activity tracking	31	25.0%
Physical activity & food tracking	20	17.0%
Period tracking	15	12.1%
Symptom | Pill tracking	13	10.5%
Pregnancy | Baby tracking	12	9.7%
Finance tracking	9	7.2%
Sleep tracking	6	4.9%
Stress | Mood tracking	4	3.2%
Water	4	3.2%
Miscellaneous	9	7.2%
Total	**123**	**100.0%**

Table 1: The number of apps appearing in each tracking domain within our corpus.

an update to both the codebook and our app corpus in the summer of 2021. In 2019, the first and last authors collaboratively conducted both deductive and inductive coding of the apps to examine both predetermined (i.e., resonance with and/or clear support for Fleck and Fitzpatrick's conceptual definitions of levels of reflection [59] via presence or absence of specific interface features) and emergent themes (i.e., additional categories of interest or any specific commonalities not previously identified in the literature) [115]. Then, in June 2021, the first and second authors revised the codebook (presented in the Table 2) based on the new app lists and more recent PI literature and re-coded the expanded corpus.

We coded *n*=123 apps based on their App Store descriptions, the screenshots included in their App Store entries, and through manual exploration of the downloaded app's user interface between June and July 2021. As part of this in-depth analysis, the first, second, and third authors downloaded and used each app in the corpus over the course of two months to ensure that we had seen (and could obtain visual evidence of) any reflection-oriented interface components present in each of the apps based on our codebook. We used each app for at least a week (up to 2–3 weeks) to collect enough data to allow for a comprehensive assessment of the apps' features properly, but did so from a perspective focused on interface exploration—identifying whether each app incorporates particular design components related to reflection—not a commitment to long-term app adoption. Rather than splitting evaluation of the apps evenly between authors, we chose a primary and a secondary coder for each category to double-check apps' features together. To choose the main two coders for each category, we considered our own personal backgrounds and interests. For instance, the first author led on the baby and pregnancy tracking apps due to recently becoming a parent. The second author's main research area is mental health, so the second author took a lead on stress and mood tracking apps. The third author is an amateur runner, so the third author focused on evaluating the physical activity tracking apps. We chose not to calculate inter-rater reliability because our coding is basically binary; whether an app has a certain feature or not [112]. Moreover, our focus in this study is to explore the possible design space of PI systems rather than report quantitative or statistical analyses. To achieve reliability, our analysis process was guided by Noble and Smith's strategies for enhancing credibility (i.e., truth value, consistency, and neutrality [127]). Specifically, the members of the research team met weekly to review the initial coding, resolve ambiguities, and identify the central themes to focus on for subsequent analysis by sharing memos that captured emergent observations. During this phase, we captured additional screenshots from each app to facilitate post hoc discussion about resolving ambiguous or uncertain coding and wrote memos to discuss new observations or possible conflicts.

5 ITERATIVE CODEBOOK DEVELOPMENT: HOW DO PI DESIGN COMPONENTS FACILITATE REFLECTION BASED ON THE PI LITERATURE?

To develop our codebook (Table 2), we combine insights from our iterative app analysis with the findings reported in the PI literature by elaborating on Fleck and Fitzpatrick's discussion of techniques for instantiating each level of reflection *in the context of PI systems*. We focus on explaining how different interface component(s) might serve as a(n) precursor for or instigator of each level of reflection in the existing taxonomy. Note that our work focuses on reflection per se, instead of its potential outcomes, although the dominant focus of PI literature is behavior change as a result of reflection.

Our coding process focused on classifying the specific reflection-oriented interface features and functionality provided by each of the apps in the corpus. We followed this approach—in contrast to directly coding for levels of reflection—for two reasons: First, coding for specific functionality (prompted by questions that we iteratively developed during our first round of coding) made for a cleaner and more definitive process than attempting to infer support for various levels of reflection directly. Second, we felt that it was important to identify and code those features that serve as the most obvious precursors to particular levels of reflection rather than coding for possible repercussions of usage (i.e., the first and most possible effect rather than the secondary effect). For instance, an individual may be able to find a pattern or correlation between different data (R2) by seeing only basic descriptions of that data (R0). However, since a straightforward recounting of the individual's data does not always afford this level of sensemaking [59], we coded instances of simple data description as R0. Likewise, we know from research that taking time to look through trends from multiple contextual data can help people construct reasonable explanations [27], but since it is not possible to infer when an interface that simply displays contextual data sequences enables this behavior, we did not code this app as (necessarily) supporting explanatory reflection (R1). Social features such as internal social networking or online community in apps may scaffold transformative reflection (R3) after users consistently interact with other users. However, we cannot make the assumption that people fundamentally change their perspective unless we observe them for any length of time. So, we focused on social features' primary effect in supporting reflection (i.e., making users describe their experiences: explanatory reflection (R1), allowing users to compare their data with others (R2)) [12, 59].

5.1 R0: Design Components for Revisiting

By definition, PI tools present data collected about individuals' behaviors and actions back to those individuals to help them monitor their behaviors, thoughts, and/or feelings. For instance, systems commonly display the number of steps recorded during a day or the duration of sleep recorded or measured. PI users check these data, especially physical activity tracking one, to look back at their recording [47]. These accounts of previous behaviors or events provide the empirical foundation for our operationalization of R0 in the context of personal informatics systems. Any representation(s) of data (e.g., textual or visualized) that describe users' status without any elaboration or explanation (i.e., interfaces that simply state a record of experiences, events, or behaviors) can support users to revisit their experiences. Systems that support goal-setting also might allow regular revisition of these personal goals by presenting the goal users set. In addition, a simple data collection feature that asks users to manually fill out data can make them look back on events or experiences [24].

5.2 R1: Design Components for Prompting and Providing Explanation

To support explanatory reflection (R1), a system can either provide a coarse interpretation of individuals' behavior or emotional states through a data-driven approach or it can prompt its users to provide an interpretation of their own data. A brief summary of recorded data can be presented alongside descriptive statistics [16], which supports a limited degree of explanatory reflection, if only by providing users with one (generated) interpretation against which reality can be compared or contrasted. This distinction between manual tracking (i.e., a user-driven approach) and automated tracking (i.e., a system-driven approach), one of the original distinctions in Li and colleagues' model of PI systems [96], also extends back into the data collection mechanism(s) implicated in R0. Some kinds of self-tracking data simply must be provided by the system's user (e.g., the subjective perception of sleep quality or mood). In other cases, PI systems can more easily capture related but not-fully-descriptive data; heart rate might relate to sleep or mood, but does not fully characterize a user's experience of these states [25]. In this sense, providing a field for manual data entry can also support R1, because as part of manual data entry, users are able to describe their rationale for or interpretation of actions. For example, Cordeiro et al. [34] found that photo-based food journaling (i.e., manually tracking food information through both photo and textual data entry) helped people to reflect on their previous experiences.

Fleck and Fitpatrick also mentioned a social aspect in PI apps can encourage users to prompt justifications or explanations [59]. An internal online community or some social networking features can be a space in which users record and share their experiences.

5.3 R2: Design Components for Comparison and Self-Diagnosis (Experimentation)

In practicing reflection level R2, individuals examine the relationship(s) between two or more data points or work to establish causality between their previous experiences and data [59]. Since PI tools are able to collect and process much information based on a wide variety of contextual factors, these systems have the potential to—and often do—facilitate these kinds of dialogic insights [26]. These representations of data from multiple sources allow people to self-diagnose their health status or to consider hypothetical questions [27, 28]. For instance, sleep tracking apps can help users understand their sleep patterns and quality of sleep through multiple contextual data with visualizations or explanations of correlation between sleep quality and other behaviors [11, 40, 98]. Furthermore, Karkar et al. introduced a research prototype that helps patients suffering from irritable bowel syndrome by allowing them to conduct self-experiments [81]. Based on this empirical evidence, in our operationalization of R2 for personal informatics, we focus on the juxtaposition of data presentation: any representation of data (e.g., texts or visualization) from more than two different temporal, spatial or other variables that enables users to consider multiple perspectives and to explore causality.

However, we should acknowledge that depending on different style of trackings, people may not gain dialogic insights from those features in PI apps [47]. For instance, for directive tracking style, temporal visualization (e.g., weekly, monthly overview) may be just a visual modality to describe their activities without elaboration (R0), it allows users with diagnostic tracking style to compare present and past achievements (R2). Bentley et al. [16] found that combining more than two temporal contextual data with a longer-term trend facilitates individuals' understanding of patterns in their behavior. Although a temporal visualization does not incorporate more than two data points, it may help diagnostic tracking users reflect on their previous experiences. This is because the diagnostic tracking is looking for a relation between more than two things [146]. On the other hand, for people who just document data without specific goals, the temporal visualization may just help them revisit their experiences unless some data violate their expectations [12, 47].

Through algorithms, PI systems also leverage users' input data to provide additional explanations with regard to an individual's experience or event [62]. These machine-driven interpretations may help users easily gain additional knowledge about their experiences or actionable insights by comparing between predictions from PI tools and current status [144].

Encountering other users' traces or patterns also provides a user with a chance to get to know an alternate perspective and to make a comparison between an individual experience and the others [13, 59, 137]. By comparing them, people become more engaged in data interpretation due to some surprise or inspiration [69]. For instance, Feustel et al. found that university students made a comparison

Coding categories	*Interface features coded*	*Possible coding values*	*Reflection level*
Data collection	*How* is data collected?	manual | semi-automatic | fully automatic	**R0/R1**
	What *type* of data does the app collect/allow collection of?	quantitative | qualitative | both	**R0/R1**
	Does the app allow users to collect data through multimedia format (e.g., pictures, video, audio)?	yes | no	**R0/R1**
	Does the app allow tracking of *explanations* or *elaborations* of data?	yes | no	**R1**
	Does the app allow collection of *flexible/arbitrary data* or markers (e.g., define their own tracking parameters)?	yes | yes, within a fixed set | no	**R3**
	Does the app include prompts to provoke users' explanations on underlying motivations or reasons?	yes | no	**R3**
Data presentation	Does the app *display collected data* (i.e., shows a record of experiences, events, or behaviors)	yes | no	**R0**
	Does the app display data collected *over time* (i.e., with a temporal dimension)?	yes | no	**R0/R2**
	Does the app show *correlation and/or causality relationships* (e.g., dashboard with two or more data shown side-by-side/overlaid, prediction based on two or more data types)?	yes | no	**R2**
	Does the app display *textual explanation(s)* about/for data?	yes | no	**R1/R2**
	Does the app allow users to customize data presentation (e.g., a customized visualization(s))	yes | no	**R3**
Goal-setting	Does the app allow setting a *personalized goal(s) (i.e., ask to choose a goal and set it up)* and displays it?	yes | no	**R0**
	Does the app allow setting a *customized goal(s) or updating/changing goal(s)*	yes | no	**R3**
Social components	Does the app show traces or patterns of other users? (e.g., statistics from anonymous data)	yes | no	**R2**
	Does the app allow users to share their data with other users through internal social networking features?	yes | no	**R1/R2**
Miscellaneous	Do facets of the app help people to consider both personal/individual and social/ethical/community issues?	yes | no	**R4**
	Does the app deliver informational or educational content?	yes | no	—
	Does the app include chatbot/avatar functionality as part of the interaction?	yes | no	—
	Does the app display quantified data (i.e., descriptive statistics or visualizations with users' qualitative or multimedia data together?	yes | no	—

Table 2: The final codebook to survey PI app features that function as a precursor of each level of reflection.

with other students who have similar backgrounds (i.e., cohorts) to identify insights and (re)frame their goals by understanding cohorts' average [55]. Other users' anonymous data can be presented in a text or visualization as well as an internal social networking feature in an app. A community can help users support not only R1 (i.e., by sharing their data) but also R2 (i.e., by checking others' data and receiving feedback from others). These social components to support reflection are also leveraged in a family [148] or work setting [111, 138]. For instance, the concept of collaborative or social reflection is used in a health and medical practice to make better decisions for treatment [111, 138]. In this context, medical workers interpret shared data collaboratively or use individual knowledge and experiences to make sense of a medical situation (e.g., treatment).

5.4 R3: Design Components for Transformation

Both Fleck and Fitzpatrick [59], and Baumer [12] argued that transformative reflection can occur with support from other dimensions (levels) of reflection. Understanding what design techniques best support transformative reflection is a more difficult and nuanced task. For instance, it is difficult to know when (or whether) people's original point of view was changed; as a result, pinpointing interface features that might provoke these kinds of changes is more difficult. Fleck and Fitzpatrick [59] also did not discuss specific HCI design techniques that support these higher levels of reflection. They acknowledged that higher levels of reflection rest upon what and how people make sense of information to become critically aware of how they perceive, understand, and feel about their experiences (i.e., internal process) [59, 114]. In the previous empirical PI studies [27, 82, 148], higher levels of reflection are rarely observed in the short-term period of interaction with PI. While it remains under-studied how to design PI to support transformative reflection, Slovak et al. proposed a framework with an emphasis on support for *the right sort of experiences* by drawing on Schön's notion of reflective practicum [155]. In the educational setting, the mentors' role is important to scaffold the right sort of experiences for mental schemas or behavior changes. To offer similar experiences with mentoring to users, PI apps can initiate a conversation about users' experiences or data to elicit interpersonal interaction [155]. Building on the concept of the right sort of experiences, in our operationalization of R3 for PI tools, we focus on a couple of features that possibly support transformative reflection. Since our work focuses on reflection, we do not refer to transformative reflection as behavior change or future action, but perspective transformation [12, 114].

Interface features that ask a question with explicit prompts is considered a helpful technique for scaffold R3 in that it helps motivate individuals to participate in data collection and encourages so-called "in-depth" reflection [95, 118, 141]. Researchers have also found that questioning (along with appropriate guidance) compels individuals to articulate rationale for their behavior, facilitating an understanding of the reasons behind actions [29, 59, 101]. Therefore, we focus on how the system prompts users to describe their behavior in a qualitative manner that facilitates a journey of making sense of experiences (i.e., right sort of experience).

Customizable approaches can also arrange the right sorts of experiences for users by eliciting their meaning making process. To reach higher levels of reflection, the main agent to actively make sense of experiences needs to be a person (e.g., a user, a student) [59, 155]. While the features we discussed in previous sections might encourage users to reassess their own perceptions, beliefs or attitude to transform their perspectives, customization features can enact agency in the reflective process to make sense of one's own experiences [5, 160]. Zhang et al. found that people are willing to make use of customization features as long as they build an appropriate level of self-efficacy [181]. Also, many first-time users often do not have a specific goal in mind when initiating PI use. However, users' goals (can) become more concrete through reflection on their behaviors collected and presented by PI tools [53, 68, 125]. In some cases, updates to a user's goal(s) can be made by using the app itself (e.g., as an outcome of experiencing the features supporting lower levels of reflection) or due to external, personal circumstances [146]. Therefore having a feature that allows users to update or customize their goals can arrange the right sort of experiences, possibly leading to R3. In this sense, customizing what data is collected, integrated, and presented can possibly help support R3 by having agency in the meaning making process [5, 87].

5.5 R4: Design Components for Transcending the Immediate Context

Finally, PI systems were originally designed to help individuals understand their own behavior(s) rather than their interactions with society or the broader world [96]. However, PI systems have since evolved to include a broader range of goals, which Lupton [109] classified into five modes of self-tracking tools: (1) private self-tracking, (2) pushed self-tracking, (3) communal self-tracking, (4) imposed self-tracking, and (5) exploited self-tracking. Among these five modes, she also identified one additional mode in which the private mode and communal mode are woven together. This 'compound' mode allows users to track their data not only to improve their lives (e.g., their well-being) but also to achieve community development or collective goals (e.g., community sustainability). For instance, *Community Mosaic* is designed to help community members improve their healthy diet practice by sharing individual behaviors [135]. While individual data is often shared with others or is visualized in a comparative context next to others' data to provide social support or recommendations [39, 122], little research has been undertaken about how best to facilitate reflection for both private and communal purposes. Reflection is also often considered as a mechanism to increase ecological awareness [20, 64], but to the best of our knowledge, PI research has focused on either individual (e.g., wellness) or ecological aspects (e.g., sustainability issues [63]) rather than the interplay of these two together.

Beyond reflection on activities, Sengers et al. argued that reflective design can support users to reconsider the impact of technology on our behaviors in a critical manner [152]. In our operationalization of R4 for PI, then, we focus on whether systems encourage users to think about ethical considerations of tracked behavior or enact collective values—interactions that help users consider not only the personal benefits of their actions, but also other entities' benefits beyond the immediate context.

6 FINDINGS: EMPIRICAL SUPPORT FOR REFLECTION IN PI APPS

In this section, we report our results about how well PI app features are provided to serve as a precursor to each level of reflection based on our codebook. Informed by a commonplace definition of personal informatics, all of the apps in our corpus collected some amount of data and provided at least minimally sophisticated descriptions or visualizations of that data to facilitate *self-monitoring*—for example, a simple snapshot of previous data or a historical record with calendars and/or time-series graphs. This follows from a commonsense assumption that design strategies in PI apps have matured and become standardized, at least to some extent. However, we also found that different levels of reflection are unevenly supported by PI apps. Here, we present details about how each level of reflection is supported (or not) through examples drawn from analysis of our corpus, and also draw attention to the instances in which control or agency over the reflection process is exercised by the system, the user, or a combination of the two (after the discussion point raised by Li et al. [96]). Figure 2 provides an overview of the design space resulting from our analysis.

6.1 Overview for Descriptive (R0) Reflection

As noted above, all apps support descriptive reflection (R0) well; all display a summary of collected self-tracking data. A calendar interface is also often used to provide an overview of users' activities for a month (e.g., *Ovia Parenting & Baby Tracker* [133],

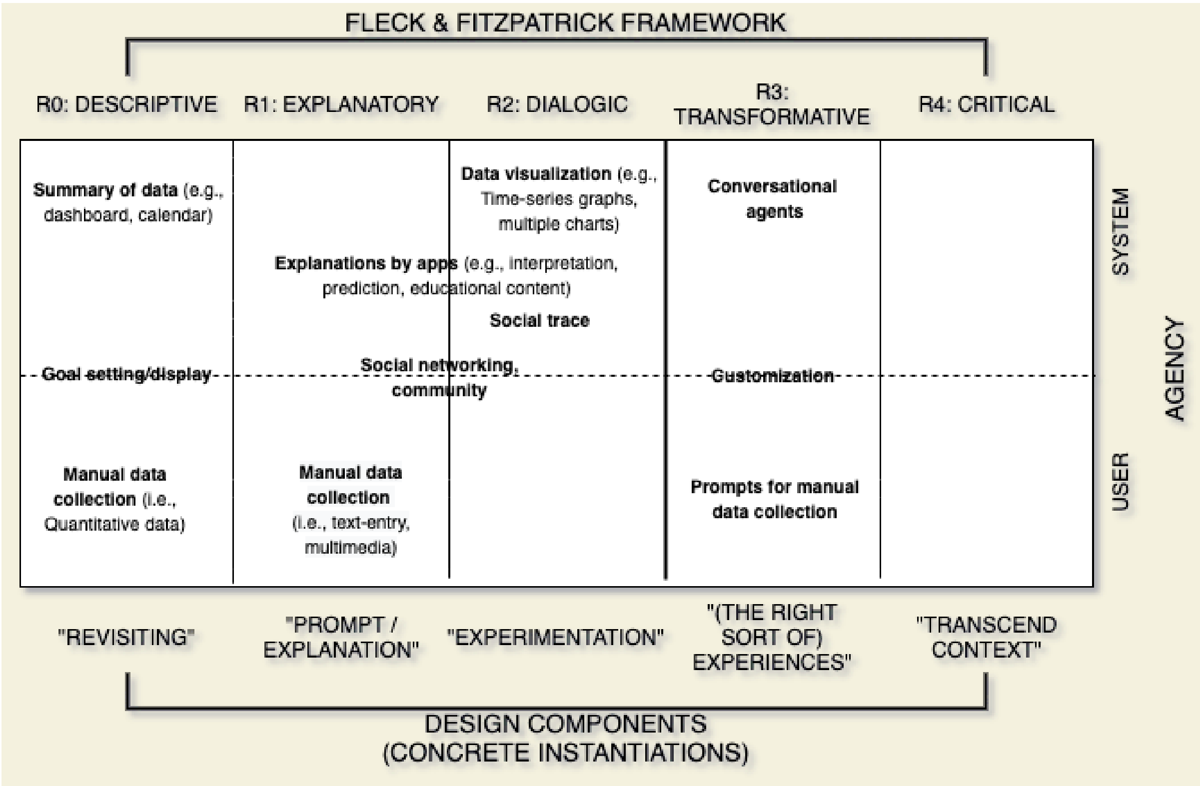


Figure 2: A design space of commercial PI apps illustrating clusters of design features organized by level of reflection [59] and locus of agency [96]

Figure 3a). The physical and food activity-tracking apps we studied show a summary of daily activities (e.g., *Argus* [7], Figure 3b). Period-tracking apps provide daily period or ovulation information, often displaying a representation of the entire current menstrual cycle (e.g., *iPeriod* [176], Figure 3c). Baby-tracking apps show a baby's recent activities (e.g., last feed time, sleep time, or diaper, e.g., *Baby Tracker* [126], Figure 3d). Apps in other categories, such as symptom or finance tracking, also presented data to help users understand their health or financial status (e.g., *EveryDollar: Budget Your Money* [164], Figure 3e). Across all PI apps, quantitative data is usually the main source for this summary type of interface.

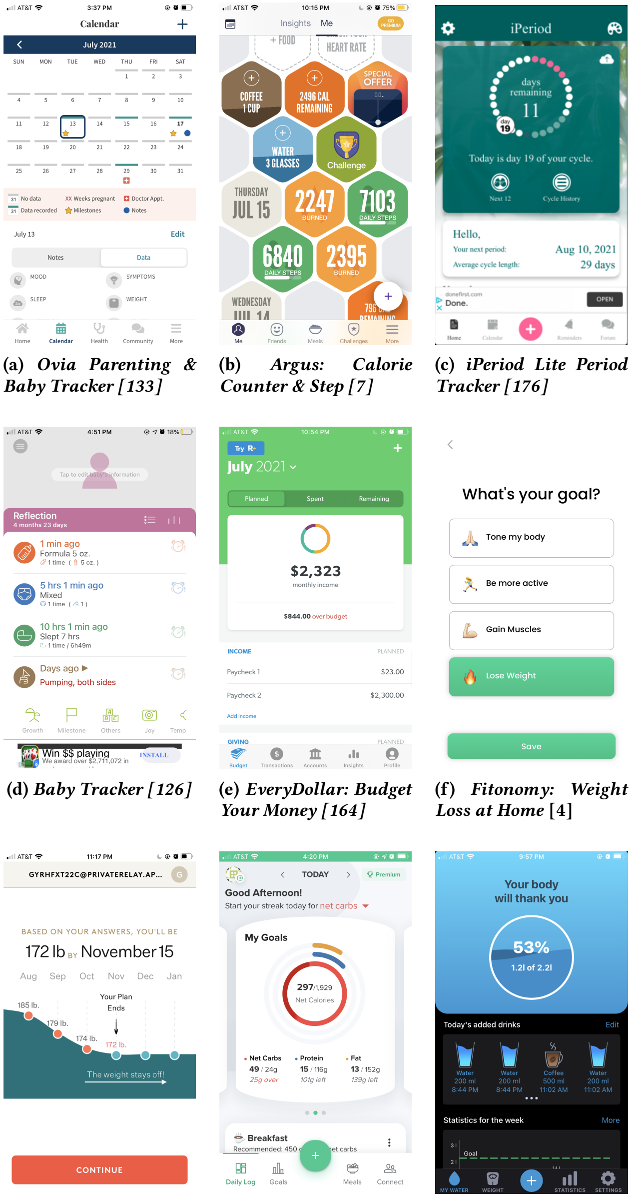
These overview interfaces (Figure 3) are often used as the default landing or "home" page in these apps, either as a mechanism for rapidly "checking in" on one's progress and/or verifying that data is, in fact, being successfully collected. Since descriptive reflection is the foundation upon which successive levels of reflection are built, the existing base of PI apps are generally well-designed to provide this base for reflection; that is, giving users an opportunity to *notice* their behavior or states [103, 170].

Most PI apps guide users in setting a goal, often by asking a series of questions during the app set-up or account sign-up phase. For instance, many weight-loss and workout apps ask users to enter basic demographic information and physiological characteristics and then use these data as the basis for establishing a personalized workout plan or calorie-consumption goal (e.g., *Fitonomy* [4], Figure 3f). Some apps even set up a target date when users should be able to meet their goal if they keep following the app's instructions (e.g., *Noom* [128], Figure 3g). These goals are often displayed in the summary interface. For instance, many wellness apps that track physical activity and food consumption highlight information about how many calories users can consume the rest of the day to stay 'on track' (e.g., *Carb Manager: Keto Diet App* [178], Figure 3h). Apps for tracking water intake also display progress reports toward daily water drinking goals (e.g., *My Water* [174], Figure 3i).

On the other hand, other types of apps such as period tracking or sleep tracking do not have a goal-setting feature. This is not surprising, because people usually make use of these apps to document their activities rather than to change behaviors [52]. Some apps related to women's health (i.e., period or pregnancy tracking) ask users to choose a main reason/goal why they use the app (e.g., trying to get pregnant, tracking menstrual cycles), but this information is usually used to set up the main focus of the user interface.

6.2 Data Visualization for Descriptive (R0) and Dialogic (R2) Reflection

Although some apps display only a snapshot of previous behavior, most (87% of apps in the corpus) help users to compare or contrast



(a) *Ovia Parenting & Baby Tracker* [133] (b) *Argus: Calorie Counter & Step* [7] (c) *iPeriod Lite Period Tracker* [176]

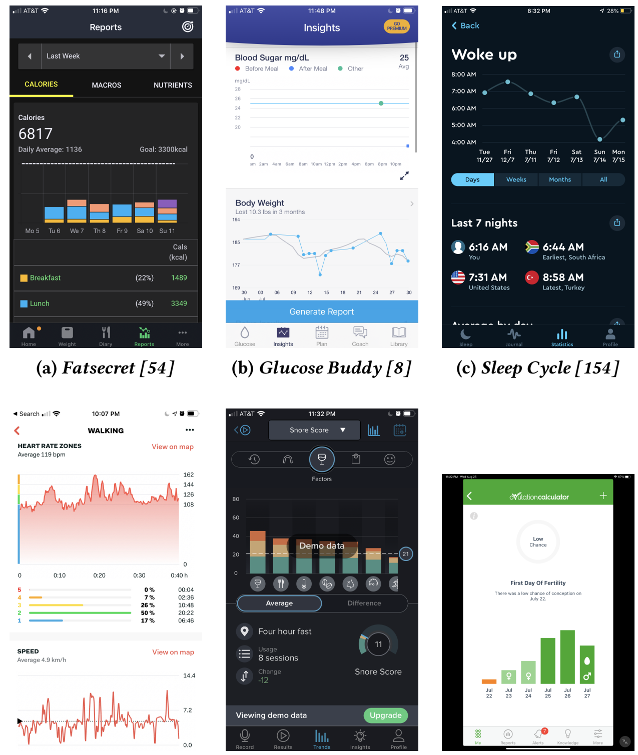
(d) *Baby Tracker* [126] (e) *EveryDollar: Budget Your Money* [164] (f) *Fitonomy: Weight Loss at Home* [4]

(g) *Noom* [128] (h) *Carb Manager: Keto Diet App* [178] (i) *My Water* [174]

Figure 3: Features that support descriptive reflection (R0) in apps.

their data by providing longitudinal, time-series graphs, overlaid or annotated with a variety of contextual factors (Figure 4). These visualizations enable users to browse their accumulated data, which is helpful for facilitating remembering or reminiscing about previous behavior [47]. These analytical representations transcend simple descriptive reflection and begin to suggest more explicit support for dialogic reflection (R2).

The most frequent visualization is to incorporate users' historical data with time on the x-axis to show temporal patterns. For instance, a common graph shown in physical, food, symptom, and pregnancy tracking apps shows changes in weight over a period of time (e.g., *Glucose Buddy Diabetes Tracker* [8], Figure 4b), since weight control is one of the main themes/goals in those PI apps. In addition to weight, temporal graphs are also used to show changes over time in a wide variety of other behaviors (Figure 4).



(a) *Fatsecret* [54] (b) *Glucose Buddy* [8] (c) *Sleep Cycle* [154]

(d) *Sports Tracker for All Sports* [2] (e) *Snore Lab* [156] (f) *Ovulation Calculator Fertile Tracker & Calendar OC* [131]

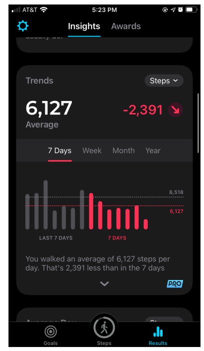
Figure 4: Features that support descriptive (R0) and dialogic (R2) reflection in apps.

To help people gain knowledge about relationships in their historical data, many PI apps display multiple visualizations together. For instance, fitness tracking apps combine display of running pace, elevation, heart rate, and weather, enabling users to identify patterns or trends from their workout history (e.g., *Sports Tracker for All Sports* [2], Figure 4d). Similarly, food tracking apps (e.g., *Fatsecret* [54], Figure 4a) displayed calorie intake, activities, water consumption, and weight visualizations together to foster exploration of how a user's overall lifestyle influences their weight. Many period-tracking apps provide not only a record of reported menstruation data, but also provide predictions of ovulation and period arrival over the course of subsequent menstrual periods in the same visualization (e.g., *Kindara* [89]). Many sleep tracking apps displayed the relationship between sleep quality and pre-sleep actions, thereby allowing users to conduct self-experimentation (e.g., *Snore Lab* [156], Figure 4f).

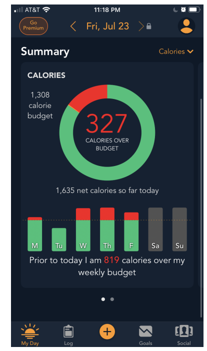
Some PI apps offer predictions to help users take action; for instance, anticipated future weight based on current food consumption and physical activity norms, the best moment for a baby to sleep, the date of and symptoms expected in the next menstrual cycle, or the chance of conception on a given day (e.g., *Ovulation Calculator Fertile Tracker & Calendar OC* [131], Figure 4g). These prediction features may allow users to project their future-self, leading to dialogic reflection on current behaviors [144].

A reflective explanation to interpret previous activities (R1) is often a prerequisite to understanding the relationships among behaviors or events (R2) in the reflective practice [59]. This is because a process of justification for action allows people to explore their experiences and their meaning(s). However, since PI apps readily provide information about correlation (and potential causality) among data, these systems aim to help individuals to engage in dialogic reflection, even in the absence of narrative explanations. These technologies greatly reduce the barriers to exploring these kinds of relations, and we can infer that this is one reason for the success and popularity of this class of applications.

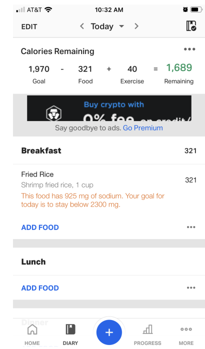
Current apps consistently provided opportunities for users to scaffold dialogic reflection (R2). However, depending on their level of motivation for achieving their goal or their degree of engagement with the app, users may or may not be able to extract insights from these relational data presentations (e.g., inferring the possible cause of insomnia). Our findings also confirmed that most current PI apps are designed based on the self-improvement hypothesis [84].



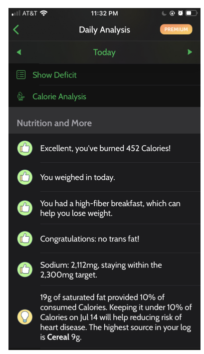
(a) *StepApp Pedometer* [158]



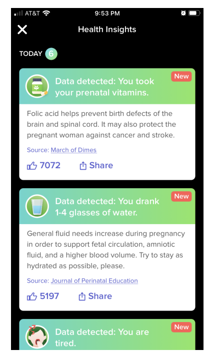
(b) *Lose it* [57]



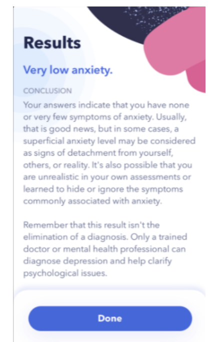
(c) *MyFitnessPal* [123]



(d) *My Net Diary* [124]



(e) *Nurture* [65]



(f) *Breeze* [10]

Figure 5: Features that support Explanatory (R1) and dialogic (R2) reflection in apps.

6.3 Machine-Driven Explanation for Explanatory (R1) and Dialogic (R2) Reflection

To support explanatory reflection, many PI apps generated explanations based on collected data. The types of these explanations are varied, including both (1) simple annotations and (2) interpretations of input data.

Some apps provide brief text describing visualizations. These explanations may not be necessary to support any level of reflection; rather, they help users understand how data are displayed. For instance, *StepApp Pedometer* [158] (Figure 5a) and *Lose it* [57] (Figure 5b) include a very quick summary of a user's data below a visualization (Another example is from Figure 4g). These annotations may help visualization novices or individuals with limited information literacy get information from graphs [70].

Other apps provide rich interpretations of users' behavior, which might otherwise have been difficult for a user to perceive without expending intentional effort, committing to a long-term perspective, and/or acquiring new scientific knowledge. By leveraging users' data, these apps attempt to provide constructive explanations in support of both explanatory (R1) and dialogic reflection (R2). For instance, *MyFitnessPal* [123] (Figure 5c) and *MyNetDiary* [124] (Figure 5d) showed short explanations of the nutritional values of foods a user has eaten and provides recommendations based on these data, perhaps to better situate each meal in terms of users' nutritional goal(s). Some stress and mood-tracking apps (e.g., Figure 5f) are based around a survey feature to diagnose emotional status. Based on users' responses, they provide interpretations to help users understand their emotional state.

We observed many instances in which apps provide informational or educational content alongside users' data or goals. These apps contextualize users' data by presenting snippets from related scientific articles or providing normalizing data ranges, with both kinds of information acting as pedagogical tools [110]. For example, *Nurture* [65] (Figure 5e) displays informative excerpts (e.g., information about prenatal vitamins) after a user enters a self-report. Although some apps post educational content in response to input data to help users interpret their behavior or experiences, in most cases, PI apps often include generalist content that might not be related to a particular user's needs or circumstances. For instance, many apps provide "How-to" instructions or tips (e.g., running or workout guide for beginners, recipes or meal plans, making a budget). App descriptions on the App Store or apps' websites often promote these informative resources as a benefit to users. However, these resources are a bit different than justifications provided by users about their own experiences. While these resources may help users to gain more knowledge about how their behavior can influence well-being, the objective descriptions offered by these systems are less likely to trigger reflection on one's own behavior. Overall, most current commercial apps focus more heavily on providing related scientific, medical, or lifestyle knowledge than in scaffolding self-reflection based on users' own data.

6.4 Data Collection for Descriptive (R0) and Explanatory (R1) Reflection

It is rare to find apps in the App store based on fully-automated data collection, except for physical tracking apps with a focus on counting steps. About half of the apps in the corpus leveraged only users' manual inputs, while the other half collected a combination of self-report data and smartphone sensor data. Some apps also provide options of connecting external devices to automatically collect data for which no mobile phone sensor currently exists (e.g., heart rate from a dedicated heart rate monitor).

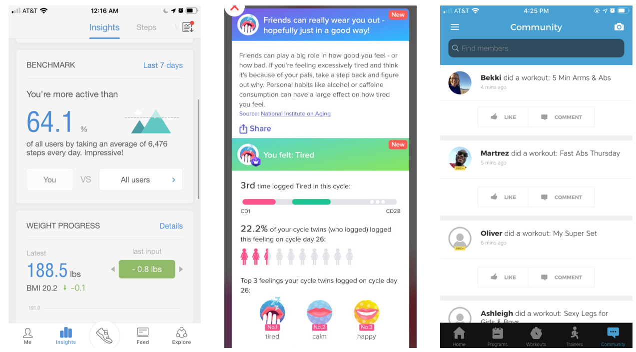
Beyond quantitative data, the majority of apps (n=79) collected other data types such as text or multimedia. This kind of open-ended text entry allows users to describe or annotate their behavior or simply capture their thoughts about a particular event (Figure 7). However, in most cases, completion of these "note" fields is optional (e.g.,*Huckleberry: Baby & Child* [80], Figure 7a). Given that PI tools emerged alongside the "*Quantified* Self" movement [177], it stands to reason that the dominant perspective embodied by these tools is that quantitative assessment can facilitate behavior change. It is likely also significant that, in most cases, quantitative data is easier to automatically sense, report, and summarize than open-ended, textual or multimedia data. For instance, no apps in our sample provided any analysis or summary of text in users' notes—users were not directly prompted to reflect on their textual data.

Among apps that did support multimedia data collection, photographs were the dominant medium. While some food tracking apps leverage users' photos to easily collect data, other apps allow users to upload photos to help remember their experiences or track bodily changes. PI tools have been often criticized for their restrictive ways of capturing how users experience and remember their lives [47]. Although PI apps, in general, put more emphasis on quantitative data, some do offer an opportunity to users to capture their lived experiences through texts or photos.

6.5 Social Components for Explanatory (R1) and Dialogic (R2) Reflection

While the social aspect of tracking has been discussed in the literature, most PI apps (n=80) in our corpus did not include social features. Community features in current PI apps appear to be focused on increasing users' self-motivation for achieving goals by sharing progress or normalizing their data against reports submitted by others with similar goals rather than collaboratively making sense of a particular shared experiences or tasks. Many physical, food, period, or baby tracking apps connect to internal (app-specific) social networks, but this is not common across other categories of apps in our corpus. For instance, many period or baby tracking apps offer connections to sub-communities depending on reported symptoms or the baby's stage of development (e.g., *Ovia Fertility & Cycle Tracker [132]*). Some physical tracking apps include sharing features that allow users to share new activities or workout histories with other app users (e.g., Figure 6d).

Social trace features that leverage other users' data to generate insights are also not common across PI apps; however, a few apps provided concise explanations to help users compare their behaviors with the other users. For example, the *Sleep Cycle* app [154] (Figure 4c) and *Pacer* [134] physical tracking app (Figure 6a) both provide



(a) *Pacer [134]* (b) *Eve [66]* (c) *Workout Trainer Fitness coach [153]*

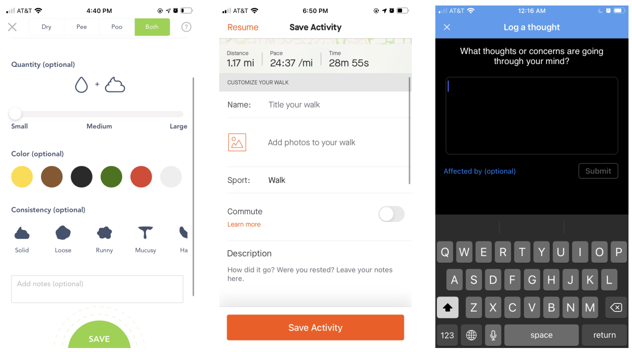
Figure 6: Social features that support explanatory (R1) and dialogic (R2) reflection in apps.

comparisons of user data with averages submitted by others. Some period-tracking apps (e.g., *Eve* [66], Figure 6b) also allow users to compare their symptoms or status with other users who are in the same menstrual cycle stage.

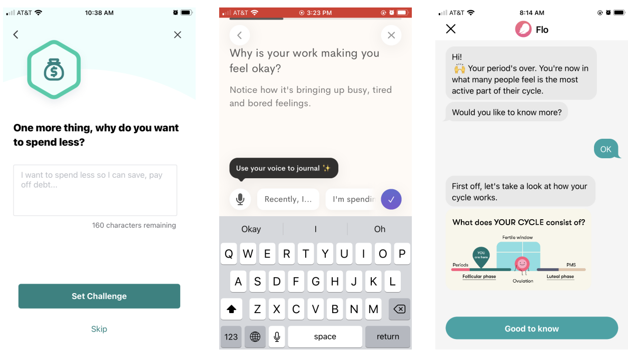
6.6 Initiating Conversations to Facilitate Transformative (R3) Reflection

We found that although many apps allow users to explain their experiences through a text-entry interface (Section 6.4), few provide any substantive stimuli to encourage users to capture their *thoughts* or *feelings* that might serve as a cue for interpreting collected data. Asking reflective questions is a key technique for promoting reflection because it encourages users to revisit their previous behavior(s) [59] and to scaffold the right sort of experiences [155]. A small number of apps (n=21) posed a question to encourage users to describe their data, including concrete prompts (e.g., *Strava* [159]: "*How did it go? Were you rested? Leave your notes here*" from Figure 7b; *RR* [143]: "*What thoughts or concerns are going through your mind?*" from Figure 7c). However, we found few apps posing *why* questions that would clearly provoke a justification for recorded behavior [95]. Given the fact that breakdowns like surprising situations are known to provoke reflection [12], a monotonous prompt (e.g., "*How was your run?*" from *Map My Run by Under Armour* [172]) is likely not sufficiently compelling to promote explanation. We did find one example of a provocative prompt in a finance tracking app: *Mint* [116]. This app asked why a user wanted to spend less during the financial planning process (Figure 7d). Some stress, mood-tracking, and journaling apps also incorporated conversational prompts (e.g., ``Why is your work making you feel okay?" from *Jour* [61], Figure 7e). By leveraging these prompts, those apps attempt to help users engage with their experiences.

Conversation is viewed as an effective way to compel behavior change in non-technology settings [91]. Since dialogue-based coaching features can scaffold transformative reflection (R3), conversational interfaces also have the potential to trigger reflection. We found that a few apps (n=10) embedded conversational interfaces into their apps. However, rather than coaching people to



(a) *Huckleberry: Baby & Child [80]* (b) *Strava [159]* (c) *RR [143]*



(d) *Mint [116]* (e) *Jour [61]* (f) *Flo [60]*

Figure 7: Features that support explanatory (R1) and transformative (R3) reflection in apps.

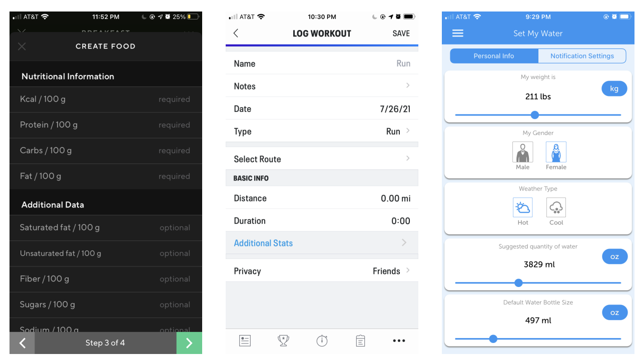
reflect on their experiences, these focused on collecting data from users or providing health-related information. In most of these cases, the process of conversation (i.e., turn-taking) was already preordained, and users are only allowed to only respond using fixed answer(s) (e.g., *Flo* [60], Figure 7f). Even though these apps provide information to users in a more natural way by adapting the conventions of a conversational interface, they fail to maximize the benefit of a true conversational grounding process in which users may be able to reflectively interpret their behaviors.

Overall, transformative reflection (R3) is poorly supported in the apps that we examined. Our findings revealed that even when open-ended text entry and conversational agent features are included, they are generally not utilized to help users to justify their behavior in support of transformative reflection. As a result, these apps may fail to capture peoples' nuanced and lived experiences, which are difficult to sense using only quantitative data.

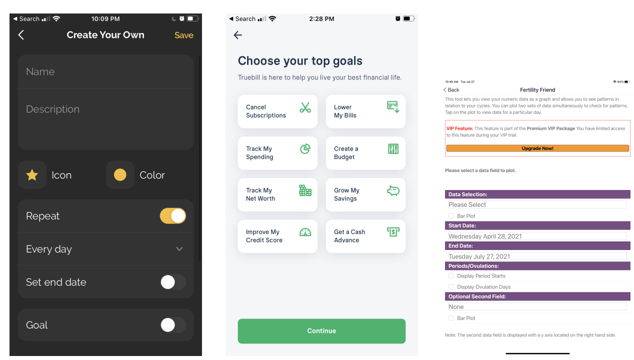
6.7 Customization for Transformative (R3) reflection

Lack of flexibility is one of the main reasons why people do not use PI tools [1]. We found that about half of the apps in our corpus provide customizable features for goal setting and data collection, but did so on a limited basis.

Various PI apps allow users to collect custom data about any of their past experiences (e.g., food, exercise, expense, emotional status, medication). For example, some food-tracking apps allow users to log meals even if a particular food has not been included in the app's database. However, most apps have a fixed list of attributes for new items; users are only able to use predefined labels to collect data. For instance, only limited information (i.e., calories burned, exercise duration) can be added to describe users' experiences completing novel exercises (e.g., *my map Fitness by Under Armour* [171], Figure 8b). Food-tracking apps often allow users to add multiple measures of nutrients and calories for new foods, but do not allow users to fully customize other ingredient information that might be personally relevant. While developing healthy eating habits or controling calorie intake are common motivations for using PI technology, people's everyday experiences in food are more diverse than these apps can capture [32, 71], perhaps due to limitations in what data can be readily visualized or algorithmically processed.



(a) *Lifesum [99]* (b) *my map Fitness by Under Armour [171]* (c) *Drink Water Reminder N Tracker [136]*



(d) *Productive [3]* (e) *Truebill [169]* (f) *Fertility Friend [162]*

Figure 8: Features that are related to transformative (R3) reflection in apps.

Some apps also allow users to explicitly create custom or personalized goals (e.g., *Drink Water Reminder N Tracker* [136], Figure 8c). In some physical work-out apps, users are able to set up a particular type of physical activity to track their exercise (e.g., walking, cycling). Some food-tracking apps allow users to control the pace of weight changes depending on their goals (e.g., losing or gaining weight); the ability to choose different food-related goals may support users' needs such as recovery from eating disorders [44]. These systems combine system- and user-driven personalization in an effort to maintain alignment with users' goals over time [73], one potential mechanism for avoiding what Epstein and colleagues refer as the "lapsing stage" [53]. For example, some physical activity- and food-tracking apps continuously update and refine goals on the users' behalf. *5k Runner* [56] adjusts the user's training plan after each run based on the measured pace and distance of the run; if a run in Week 3 went particularly well, the app might advance the user to Week 5.

However, as Munson has already identified, one limitation of these approaches is that users are only given the opportunity to select from a restricted set of goals with prefigured labels [119]. For instance, most PI apps allow users to update their goals in a quantitative manner. Although a few apps (e.g., *RR: Eating Disorder management* [143], *Sanvello:Anxiety & Depression* [149], and *Productive* [3]) provide the option to add an open-ended, customized goal, in most cases, users can only adjust their quantitative goals (e.g., calories, weight). In another example, a central feature of one financial tracking app (i.e., *Truebill Budget* [169]) is to negotiate bills with service providers and help users to cancel unwanted subscriptions. Because of this unique feature, one of the choices on this app's goal list is "*cancel subscriptions*" (Figure 8e).

Features for customizing data presentation are also not widely available in the PI apps in our corpus. Most apps only allow users to add or remove prefabricated visualizations from the main dashboard; users can customize what they want to see first, but cannot design their own visual components to capture personal contexts [87]. Across all PI apps from our corpus, *Fertility Friend FF App* [162], a women's health tracking app helping users to track period, ovulation, and fertility, has the most interactive feature to support customization. In this app, users can select one or two subset(s) of historical data (e.g., period start and end date and BBT) and choose to present their data in a line or a bar graph in a single chart with different colors.

6.8 Beyond Ego-centric components for Critical (R4) Reflection

We identified very few apps that directly support critical reflection (R4). This finding resonates with Fleck and Fitzpatrick's argument that technological support for this level of reflection is both rare and difficult to implement [59]. It appears that current PI apps primarily aim to help individuals achieve their personal goals, embodying the ego-centric perspective common in the personal informatics and quantified self domain. We did note one exception: the app *Charity Mile* [21] was designed to foreground a sense of "being" within a broader global community; its goal is to link an individual's activity (steps, tracked via pedometer) with support for NGOs.

7 DISCUSSION

Our findings contribute to a richer understanding about how the construct of reflection is currently embodied—and how it is not—in PI apps. While a certain amount of time and effort is necessarily required for reflection [59], our findings reveal that current PI tools are designed to avoid these barriers to support reflection. They make it possible for users to collect substantive amounts of data about themselves, datasets that would otherwise be unavailable without a massive time commitment to data collection, organization, and maintenance. These tools also give users access to an alternative perspective on their behaviors by providing visualizations and/or personalized explanations. Social features in some PI apps encourage users to describe their experiences and to observe other users' experiences, but they are not as widely leveraged for diverse purposes. These functionalities give contemporary PI apps the potential to foster reflection, since users may gain some knowledge through interacting with the tools. However, it is clear that the full process of reflection, based on Fleck and Fitzpatrick's framework, is not well supported in these apps. We discuss this apparent gap between reflection in theory and reflection in practice and present design considerations for better supporting reflection in PI.

7.1 Agency in Reflection: The Misnomer of Reflection and Fallacy of Insight in PI

Our findings showed that most contemporary PI apps offer lesser agency to the individual user for reflecting on their data. Since systems pre-constrain the meaning-making process, there is less opportunity for individuals to generate knowledge beyond the information displayed by the system. Boud et al. [18] emphasized the role that people themselves play in reflective practices in the context of learning. Even though teachers can intervene with various techniques to encourage and support learners in reflecting on their experiences, there is a limited effect on stimulating reflection without learners employing their own meaning-making process. Following only predetermined sequences can limit users' reflective thinking [129]. Hatano and Inagaki also warned of the risk that "our modern technology, which aims at reducing built-in randomness in the system, by no means facilitates the acquisition of conceptual knowledge" [77]—that is, streamlining the reflection process by placing a heavier reliance on system-driven reflection support results in fewer opportunities for users to acquire self-insights and reflect on their own experiences. Even though reflective informatics is typically considered an intervention empowering users to self-experiment in service of their own wellness [13], our findings suggest it might be just another mode of *persuasive technology* [20], designed with a restrictive perspective of behavior changes facilitated (e.g., monistic, fragmented) [142]. Since system-driven reflection is a predominant design strategy in PI apps from our corpus, they may encourage people to automate behavior change through information and instructions *prescribed by others* rather than supporting a *generative* meaning-making process by individual users. This is one possible reason why most reflective practices supported by PI systems remain in the lower levels of Fleck and Fitzpatrick's [59] taxonomy of reflection. Kersten-van Dijk et al. also found that PI literature that rests on the self-improvement hypothesis for the prevailing intention in designing PI systems and acknowledges the difficulty of capturing the process from self-awareness to transformation [84].

We argue that current design strategies in PI apps ignore one important facet of reflection—reflection as a conversation between the subject and the situation that enact the meaning-making process [30, 151]. Donald Schön introduces the idea of reflective practices (i.e., reflection-in-action) with an emphasis on the individuals' conversation with the materials of their practices [151]. In contrast, PI apps make it convenient to track data without ever supporting the user in considering the question of *why*, an essential facet in scaffolding reflection. The PI model [96] affirms that reflection has occurred as long as people simply look at the data. So PI users are usually exposed to a representation of self or world that is mainly understood and processed by a system. While the notion of reflection-in-action provides an explanation about how PI helps people gain insights, especially through "glanceable" information visualizations (e.g., [137, 145]), the value of reflective practice—a user's appreciation of the situation or experiences to question the assumptional structure of knowing-in-action as a way of reframing problems [150]—may be overlooked the design of PI. Schön highlights the value of reflective practices in response to technical rationality that is guided by engineering-type problem-solving approaches [151]. Rather than making use of knowledge from formal learning or guidelines, Schön argues that people are able to resolve the situation and develop new knowledge by reflecting on a situation/task/issue in the midst of action (i.e., an alternative epistemology of practice). So, reflection is not supported by repetitively following a prescriptive process; it emerges from a practice of noticing different things and exploring alternative ways of seeing one's experiences and behavior [151]. Therefore, our findings urge that the self-improvement hypothesis [84] needs to be reconsidered; leaning on only the system-driven approach for supporting reflection limits effective transformation of self-knowledge into new behavior or mental schemas and limits the consideration of wider implications beyond the self.

Consider, for example, the many PI apps designed to support personal wellness. The lack of consistent support for user-driven reflection in these apps could have unintended consequences that undermine their value proposition. For instance, by highlighting only particular and isolated information such as consumed calories in a numerical (quantitative) fashion, the design of current weight loss apps have been shown to aggravate eating disorders [44]. Kersten-van Dijk et al. also found that a system's data visualization based on standardized rules such as "10,000 steps a day" may lead to misinterpretation of data and self [173]. These misuses of PI apps could potentially be overcome by encouraging reflection on previous use or on users' holistic health journeys; however, the lack of support for reflective practice in these systems can lead to negative impacts on personal health—and make these negative impacts last longer [44]. Our findings point out that the system-driven reflection supported by current PI apps could be one point of failure for helping individuals make sense of their everyday experiences. Although PI tools certainly provide new information (e.g., information visualization, informational or educational content) to help people gain insights—often, data that supports the awareness of the current status and the construction of correlational or causal narratives—the absence of support for user-driven reflection in PI tools may lead to impoverished or incomplete insights, since these design decisions deprive users a key part of the meaning-making process. Even though the lived informatics paradigm places an emphasis on how tracking is experienced in people's everyday lives [146], these apps still by and large rely on machine-assisted routinization (i.e., reinforcing a drumbeat of evaluating a daily numerical score), revealing a potential blind spot in current models of PI design and use.

We do not argue that system-driven approaches to supporting reflection do not contribute to scaffolding reflection. Rather, user-driven approaches also should be part of design strategies in PI to help users generate organic insights instead of prescriptive ones based on their lived experiences. To better stimulate reflection, PI apps must share agency in the meaning making process with users.

7.2 Design Implications: Opportunities for Better Supporting Reflection in PI Apps

How might we better design PI apps in order to encourage people to doubt their original assumptions, contemplate the mechanisms embodied by these systems, and critique the representations that they provide? What type of technological intervention would be most beneficial in supporting multiple—or perhaps all—levels of reflection?

Our findings show that depending on which app a user selects to track their data, the app selection decision preconfigures which data are collected and presented for reflection. In other words, selecting a certain app forces users to commit to a restricted boundary of reflective practice; current systems limit people from having individually-driven insights based on human agency. Drawing on mixed-agency approaches such as the notion of co-performance [93], we present design implications that PI app designers may consider in order to help users engage with their experiences. Based on tracking styles and goals, not all levels of reflection need to be supported [52, 68, 146]. So, our goal is not to generalize one-size-fits-all design implications for all PI apps. Rather, we discuss potential opportunities for innovation in the design process of PI apps—ways to think about overcoming blind spots of current apps in supporting meaning-making based on lived experiences. To do this, we draw on PI models [53, 96], focusing on the *collection*, *integration*, and *presentation* stages[3].

7.2.1 *Data Collection: Moving Beyond the Quantified-self and Toward the Qualitative-self.*

Most PI tools embody the notion of "self-knowledge through number," widely known as the *Quantified Self* (QS) movement [177]. However, quantifying behavior can lead to assumptions about what it means to be "normal," healthy, or beautiful, and this can result in feelings of inadequacy and even anxiety if one does not fit the standard [72]. While there are beliefs that personal data can be objective, some researchers have raised the concern that PI tools are inherently biased and political [72, 102, 157]. We do not argue against the importance of data accuracy in PI systems. However, what if quantified measurements of a person's health are not compatible with that person's lived experience, even if it is scientifically correct?

Since most data processing—from collection to representation—is quantitative and/or numeric, our lives are also interpreted numerically in most PI apps. Elsden et al. [47] argue that how PI tools make sense of our lives is quite different from how individuals do. For instance, while knowing how long we've run and how much our

[3]The fourth stage of Li et al.'s model is technically labelled "reflection"; we re-label it here to better distinguish it from our more expansive discussion of reflection.

Design implications	Current PI apps' dominant design approaches	Considerations
Move beyond the quantified-self and toward the qualitative-self	Primarily collect, interpret, and present data in quantitative format	• Leverage (1) provocative prompts (i.e., first-person narration techniques [141]) to help users make sense of their experiences in qualitative manners, and (2) smart journaling feature to help users record personal thoughts or feelings. • Consider how and what system driven features can be combined with qualitative-self approach for supporting meaning-making process • Consider users' tracking goals and motivations to decide whether manual tracking is effective • Consider novel interfaces such as conversational interfaces as a way of tracking/collecting qualitative data to reduce users' burden in manual tracking
Empower users through customizable design	Less controllability in customizing data processing	• Consider more customizable design features that empower users by allowing them how they want to be represented and what they want to achieve based on their lived experiences • Consider possible external/physical materials that can support customization on data presentation
Data presentation based on mixed agency	Focus on presenting only accurate or relevant information	• Consider displaying a combination of system-driven (e.g., visualization based on quantitative data) and user-driven information (e.g., journaling)

Table 3: Summary of design implications

performance has been improved can be helpful in understanding running behavior, remembering how an individual *felt during* and *thought about* a workout after the fact are important in gaining an alternative perspective and evaluating goals [165]. Although capturing these feelings is important for fostering reflection [59], the value of these *qualitative* data is currently underestimated in the design of contemporary PI apps. It is not common to incorporate interventions that prompt open-ended responses in which people rethink their underlying motivations or reasons. Instead, most PI apps focused on collecting numerical or ordinal data that can be measured in a quantitative manner. Swan [161] notes that tracking of qualitative phenomena are important when seeking to understand subjective data because qualitative data more effectively connotes nuance and complex facets of people's experience [48, 87, 141]. People also usually express their goals in a qualitative manner at the beginning of PI use [125].

A qualitative-focused approach can function as an impetus to support a sense-making practice based on noticing behavioral differences over the long term, resulting in changes to mental states or behavior [142]. Rapp and Tirassa argue that supporting subjective interpretation of personal data by leveraging first-person narration techniques (i.e., why and how questions) is crucial for making PI use meaningful [141]. Similar to previous work [95, 141], we also suggest more future research on how best to incorporate persuasive and provocative prompts into the design of PI tools is needed to promote users to describe the rationale for or interpretation of actions or events. This approach can more effectively scaffold reflection and generation of *self*-insights because answering provocative questions encourages people to engage in interpretation of their experiences or behaviors by themselves, a key precursor to effective self-reflection. For instance, by prompting users to describe why they want to set a certain goal (e.g., Figure 7d), rather than only asking them to set a quantitative goal, users may create more meaningful goals derived from their lived experiences [125].

Some research suggests that a user-driven, qualitative approach such as journaling (e.g., [6, 63]) or narrative capture tools (e.g., [79, 121]) can be a means to improve reflective practices. The journaling and mood tracking apps in our corpus leverage prompts to scaffold reflection relatively well (examples from Section 6.6). When writing in a diary to record personal thoughts and feelings, people are able to revisit and reevaluate their previous behavior and to consider the wider implications of those behaviors (i.e., reflection-through-recording) [63]. By accounting for everyday experiences, people who use journaling apps naturally reflect on their memories, which stimulates the meaning-making process [46]. In this process, technology functions as a "witness rather than narrator" [46, p.2827] to capture experiences. Elsden et al. however, found that system-driven features of remembering (i.e., recorded contextual data such as location or weather) provide authenticity for user-driven reflection (i.e., journaling). So how mixed-agency in journaling features can be designed to stimulate meaning-making is an interesting avenue for future research.

While the benefit of journaling practice in supporting reflection might be obvious, employing a manual user input mode comes with tradeoffs; a practice of journaling might restrictively support a particular style of tracking (i.e., documentary tracking [146]) rather than a breadth of goal-oriented tracking approaches [45]. So, PI designers should consider people's tracking types and their goals when applying those features. Moreover, manual tracking puts a burden on users, resulting in the abandonment [50, 181]. Therefore, designers should consider that the qualitative-focused approach is based on people's motivations; people with less intrinsic motivation might not naturally or willingly engage in recording their activities or thoughts.

To reduce users' burden in manually tracking these valuable qualitative data, the potential of multimodal data collection and conversational interfaces might be more prominently considered in the design of PI apps. Research has shown that voice interaction can reduce the data collection burden when users have difficulty in manually inputting their data [106]. Further, conversational interfaces have been shown to facilitate journaling practices and elicit reflection on previous behaviors in order to provoke critical reflection and action [90, 91, 147]. We found that applications of conversational interfaces such as chatbots are used in very limited numbers of popular PI apps (see section 6.6), and in most cases, they focus on delivering information. We suggest that future research investigate how provocative prompts to elicit justifications for reflection can be effectively applied in PI tools in consideration of the potential of multimodal and conversational interfaces.

7.2.2 Data Integration: Empowering Users through Customizable Design. Our findings suggest that allowing users to participate in not only the data collection phase but also the data integration phase—where they are actually able to customize the meaning-making process—is an important way to facilitate reflection by entangling PI use with personal contexts and lived experience. Khovanskaya et al. [86] emphasized that PI systems' internal mechanisms—"decisions about which data are collected and how data are reflected back to the user" [p. 3411]—should be rendered transparent for users in order to provoke their critical reflection on the effects and outcomes of these systems' use. Even if a PI system does not provide full controllability of the data integration process by its users, disclosure about how a system's suggestions (i.e., goal-setting) are generated can help to foster users' commitment to their goals [180]. Because the types of data collected and integration algorithms employed are predetermined by app developers, there is an acknowledged need to design these apps to be more intelligible by informing users about "what they (PI apps) know, how they know it, and what they are doing with that information" [15].

However, more can be accomplished beyond improving the intelligibility of these apps' internal mechanisms. To help PI users find meaning in their everyday lives, the need to pay attention to the customizable design of PI has been discussed previously (e.g., [140, 173]). Since reflection is interwoven with lived experience, a degree of flexibility to actually *control* data processing (i.e., collecting, managing, and displaying data) in PI tools would be a valuable improvement for supporting reflective practices. There have been efforts to explore how to develop systems that empower users by allowing for data customization (e.g., [87, 88]). However, our findings provide empirical evidence that customizable features in PI apps—(1) the variables and types of data that the system collects and how to visualize them, and (2) their goals for using the PI tool to better support reflection—are still uncommon. *Intended* self-tracking practices by app designers are often aligned with individuals' existing goals or motives [146]. Currently, these goals are interpreted by the preconfigured mechanism(s) provided by the apps and aligned with the system's features to encourage people to track their data. However, the lack of customizable design causes abandonment of PI systems [94]. Previous work argued that PI tools should be designed to promote a better experience of and in our lives, taking into account users' expertise in a tracking domain rather than promoting fixation on the tools and the tracked data [48, 140]. To promote reflection on individuals' personal lives, the importance of customizable design should be underscored in the design of PI apps.

We found that some current apps provided good support for customizing the tracking of any data type that users would like to track. However, merely providing more opportunities for users to collect diverse data does not necessarily mean that people can customize data collection and exploration to meet their needs. When some instances of this kind of flexibility were observed in contemporary PI apps, these systems still made use of predetermined data-processing pipelines.

Allowing users to choose to display only a subset of data or visualizations can be helpful in order to avoid information overload. Our findings showed that a few apps allow users to customize their dashboard, but there was no feature that supports users to encode their experiences to visualization by designing personalized visuals. Ayobi et al. presented the potential of how physical tracking practices can be converted into digital technology [5]. It is known that constructing visualizations through physical materials is effective in scaffolding not only lower but also higher levels of reflection [130, 167]. For instance, Khot et al. showed that materializing physical activity into a physical artifact facilitates reflection and reminiscence on previous activities [85]. Future research should investigate to what extent and in what ways customization features for data presentation can be developed in current PI apps.

7.2.3 Data Presentation: Information based on Mixed Agency. Our study shows that current PI tools provide rich information visualizations and summaries to elicit further insights that are difficult for users to capture on their own. Combining a user-driven visualization approach with the predominant, system-driven one to provide the "right sort of experience" [155] can be useful for considering how PI apps present information to better support reflection. It is uncommon, but we found that some apps display not only data visualizations based on quantitative data annotated with users' photos or narrative textual data. Current PI apps' focus on delivering more accurate machine-driven meaning-making on individuals' data by leveraging (mainly quantitative) data. However, given the importance of individuals' meaning-making process to stimulate reflection, linking insights from machine-driven components with personal annotations such as their notes or multimedia would be beneficial to support the full process of reflection.

7.3 The Incorporation of Diverse Human Interests in Reflective Practice: Beyond Ego-Centric PI Design

PI tools might serve a valuable purpose not only for improving individual wellness but also for embodying and reflecting on cultural and societal values [152]. However, as noted previously, we identified no features in our corpus of apps that support critical reflection (R4), although one app (*Charity Run*) is designed to help people to bring global issues to mind as a motivation and context for self-tracking.

The limited perspective of human interests in discovering knowledge and reflecting on their experiences may cause the lack of support for critical reflection in the PI domain. Individuals' self-interests in pursuing knowledge and action are important determinants of the way(s) in which an individual is reflecting. Habermas [74] viewed reflection as a tool used in the development of particular forms of knowledge, depending on people's interests [83, 113, 175]. He introduced three different interests: (1) technical, (2) practical, and (3) emancipatory. These interests determine the mode of making sense of experiences and discovering knowledge. Most technologies are often designed to enhance instrumental knowledge to help people have better control over their situations, building on Dewey's notion of reflection as a problem-solving process to cultivate instrumental knowledge [42]. Drawing on this notion, PI apps are also designed to cultivate self-knowledge to control our physical body, affective aspects, or behaviors by aligning a personal goal. So the main discussion to support reflective practices in PI use focuses on how to cultivate instrumental knowledge.

We argue that the practical and emancipatory interest should be also considered in designing PI apps to support critical reflection. The practical interests help people to generate practical knowledge for mutual understanding and wise action within a coherent framework of values transcending an ego-centric perspective. Our review revealed that some PI apps in our corpus provided comparative data from anonymous cohorts who have a similar demographic background (e.g., *Sleep Cycle* [154] Figure 4c) or who are experiencing a similar physiological state at the same time. These PI-supported social interactions are often leveraged only to support an individual's goal or fulfill self-satisfaction with a focus on individual cognitive activity [138] rather than promote community values or collaborative work. Although personal data are derived from the individual, as long as data are collected by these systems, personal data no longer belongs only to the individual—they belong to the system. Each individual's data could be used or exploited in improving the system's integration and reflection mechanisms (e.g., recommendation algorithms) for providing information to other users. Tisné [168] argues that data should be viewed as a type of externality, similar to carbon dioxide. Although the amount of carbon dioxide emitted by a single person is scarcely detectable, the sum of carbon dioxide in our community has a tremendous effect on the environment. It may be important to support people in understanding the social reach of data sharing and toward what purpose(s) their data might be utilized in order to encourage users to reflect on its societal value. Moreover, PI app designers may consider how they fully leverage community features as not only a space for sharing and tracing others' data but also a platform able to support the whole process of collaborative reflection — (1) data documentation, (2) individual reflection, (3) collaborative reflection, and (4) sustaining outcomes [138] — in order to incorporate practical interests in PI use.

With emancipatory interests, people aim to challenge taken-for-granted assumptions to reveal the dominant values in society [83]. Critical design research is usually grounded in the emancipatory interest to make people more critical about their everyday lives by leveraging design [9]. While a critical approach to design (e.g., slow technology [75]) can be another design strategy in PI to support reflection, little attention has been paid to it in designing PI apps.

In a broad sense, we argue that designing PI tools to support not only personal wellness but also environmentally sustainable values or community-scale social determinants of health can open a new design space for PI systems. For instance, food practices (e.g., consuming food, shopping food, and disposing of food) are highly related to sustainability issues [23, 31, 63]; however, all food tracking apps in our corpus focused on personal health—specifically, weight management. This finding suggests a significant design opportunity for PI tools to transcend a limited, ego-centric perspective. The mode of tracking individual data can be also customized based on human interests. For instance, one person who is only interested in personal health issues can be overwhelmed if an app presents all information in terms of individual, society, and environment. To support not only personal wellness but also sustainability or societal issues, we suggest that PI app designers may consider embracing multiple human interests to support the full process of reflection.

7.4 Limitations and Future Work

This research has a number of limitations. Since we focus on the design elements in current PI apps rather than empirical data from actual PI users to study how PI tools support reflection, our findings cannot reveal how reflection occurs in people's everyday lives or how people interact with PI apps. Although reflection can be stimulated outside the strict confines of the apps, our findings cannot capture these reflective practices. Our study also cannot identify any secondary effect(s) of the systems' design components from a long-term perspective. To minimize this limitation, we grounded our approach in theoretical frameworks of reflection and previous empirical studies of PI. We believe that our findings reveal opportunities for future empirical work to better understand how certain PI features actually affect reflection in people's everyday lives.

Furthermore, our findings in this study are grounded in a theoretical lens derived from Western traditions of reflection. Therefore, our findings and implications might not fully encompass different values, norms, or practices that might be employed in other cultures' reflection processes. Future work should interrogate the issues described here in designing PI apps relative to local theoretical frameworks (e.g., a Daoist perspective [163]) and non-Western lived experiences [22].

Our sampling method also necessarily results in some limitations to the generalizability of our findings and implications. We aimed to select PI apps that are likely to be most widely used, but it is possible that our corpus does not represent all types of PI apps. Our sampling criteria, for example, may have filtered out some more recent apps with newer features for reflection apps that require additional wearable devices (e.g., Fitbit), or apps tailored to the needs of a very particular sub-community. However, as Kim et al. noted [88], apps in the same domain have very similar features; collecting data and the means for presenting those data are very standardized. As a result, future work could expand our framework to other apps in order to understand the impact of new, additional, or complementary features. Some apps also included design features that extended beyond the bounds of the app, itself (e.g., widgets, notification, or wearable devices); we did not explore these extra-application features in our analysis, and their impact might also be investigated in future work. Finally, some apps' advanced analysis and reporting features are available only to users with paid subscriptions, limiting our ability to exhaustively examine the complete feature set of some PI apps.

Our findings showed that depending on an apps' domain, design features that support reflective practices are quite different. These findings inform further research on how reflective practices look in a particular domain and in what way design features can be utilized to support a specific user's purpose. The goal of this study was to understand how reflection is supported by a number of different design features in general, rather than in a particular domain, such as physical activity tracking. We hope our findings can serve as a baseline or starting point for investigating domain-specific PI cases.

Lastly, our findings raise a fundamental question: *do people need to reflect on their experience to achieve their goal (i.e., to affect a behavior change) at all?* Our findings reveal that PI apps do not, in general, foster the full process of reflection well, but they are still very popular in the commercial marketplace. The underlying assumption in our study is that reflection is beneficial for people to gain new knowledge, to reconstruct mental schemas, or to induce behavior changes for positive outcomes. Most literature in the PI domain has an embedded assumption that the PI system should always be driving users to reflect on their data. However, in some cases, reflection does not need to be an indispensable component in PI use. This is because reflection sometimes leads to an unexpected or negative outcome [12, 36, 43, 118]. Depending on why a PI tool is being used, the appropriate or desired mode of reflection might be different, as well. For instance, Gouveia et al. [68] found that most activity tracker PI users just glance at their tools ("brief, 5-sec sessions where users called the app to check their current activity levels with no further interaction") rather than meaningfully or deeply engaging with their data. In this case, stimulating higher levels of reflection might not be necessary; supporting the lower levels consistently and well might be sufficient. Future research should investigate the relationship between users' goals for the use of PI tools and the varying levels of reflective practices to explore whether (and how) reflection plays different roles in service of different goals.

8 CONCLUSION

In this research, we developed a theoretically and empirically informed codebook and design space for how reflection might be instantiated in PI apps, and we surveyed commercially available PI apps to understand the extent to which they actually support different levels of reflection. In total, we identified and characterized the gaps between theoretical research on reflection and interface features in current apps to suggest a number of ways in which PI tools often fail in supporting reflection, particularly user-driven reflection. Models of PI systems [53, 96] have contributed to improvements in features that increase engagement in tracking, avoid the abandonment of tools, and help people achieve their goals. Although these are all important goals in system design, it is also necessary to consider how to design PI tools to better promote the conscious awareness of and reflection on previous behaviors or events. In designing the next generation of PI tools, then, we advocate for prioritizing helping people engage in meaning-making to cultivate their own insight and, in doing so, help to make individuals' everyday lives more meaningful.

ACKNOWLEDGMENTS

We thank the reviewers for constructive feedback that helped improve this paper. We also thank Geraldine Fitzpatrick and Amy Voida who provide constructive feedback on this work. We are also grateful to Junnan Yu, Fuji Robledo Yamamoto, and Shiva Darian for their feedback on drafts of this work.

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. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 115 (Dec. 2020), 24 pages. https://doi.org/10.1145/3432236

[2] Amer Sports Digital Services Oy. 2021. Sports Tracker for All Sports. Apple App Store. https://apps.apple.com/us/app/sports-tracker-for-all-sports/id426684873

[3] Apalon Apps. 2021. Productive - Habit Tracker. Apple App Store. https://apps.apple.com/us/app/productive-habit-tracker/id983826477

[4] Appostafat GMBH. 2021. Fitonomy - Personal Trainer. Apple App Store. https://apps.apple.com/us/app/fitonomy-weight-loss-at-home/id1107242759

[5] Amid Ayobi, Paul Marshall, and Anna L. Cox. 2020. *Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care.* Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376809

[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.* ACM, New York, 1–14.

[7] Azumio Inc. 2021. Argus: Calorie Counter & Step. Apple App Store. https://apps.apple.com/us/app/argus-calorie-counter-step/id624329444

[8] Azumio Inc. 2021. Glucose Buddy Diabetes Tracker. Apple App Store. https://apps.apple.com/us/app/glucose-buddy-diabetes-tracker/id294754639

[9] Jeffrey Bardzell and Shaowen Bardzell. 2013. *What is "Critical" about Critical Design?* Association for Computing Machinery, New York, NY, USA, 3297–3306. https://doi.org/10.1145/2470654.2466451

[10] Basenji Apps. 2021. Breeze: mood tracker, diary. Apple App Store. https://apps.apple.com/us/app/breeze-mood-tracker-diary/id1450365119

[11] Jared S. Bauer, Sunny Consolvo, Benjamin Greenstein, Jonathan Schooler, Eric Wu, Nathaniel F. Watson, and Julie Kientz. 2012. *ShutEye: Encouraging Awareness of Healthy Sleep Recommendations with a Mobile, Peripheral Display.* Association for Computing Machinery, New York, NY, USA, 1401–1410. https://doi.org/10.1145/2207676.2208600

[12] Eric PS Baumer. 2015. Reflective informatics: conceptual dimensions for designing technologies of reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM, New York, 585–594.

[13] Eric P.S. Baumer, Sherri Jean Katz, Jill E. Freeman, Phil Adams, Amy L. Gonzales, John Pollak, Daniela Retelny, Jeff Niederdeppe, Christine M. Olson, and Geri K. Gay. 2012. Prescriptive Persuasion and Open-Ended Social Awareness:

Expanding the Design Space of Mobile Health. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (Seattle, Washington, USA) *(CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 475–484. https://doi.org/10.1145/2145204.2145279

[14] Eric PS Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing reflection: on the use of reflection in interactive system design. In *Proceedings of the 2014 conference on Designing interactive systems*. ACM, New York, 93–102.

[15] Victoria Bellotti and Keith Edwards. 2001. Intelligibility and accountability: human considerations in context-aware systems. *Human–Computer Interaction* 16, 2-4 (2001), 193–212.

[16] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 5 (2013), 30.

[17] Marit Bentvelzen, Jasmin Niess, Mikołaj P. Woundefinedniak, and Paweł W. Woundefinedniak. 2021. The Development and Validation of the Technology-Supported Reflection Inventory. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 366, 8 pages. https://doi.org/10.1145/3411764.3445673

[18] David Boud, Rosemary Keogh, and David Walker. 2013. *Reflection: Turning experience into learning*. RoutledgeFalmer, London.

[19] Dionne Bowie-DaBreo, Sandra I Sünram-Lea, Corina Sas, and Heather Iles-Smith. 2020. Evaluation of treatment descriptions and alignment with clinical guidance of apps for depression on app stores: Systematic search and content analysis. *JMIR Formative Resesearch* 4, 11 (13 Nov 2020), e14988. https://doi.org/10.2196/14988

[20] Hronn Brynjarsdottir, Maria Håkansson, James Pierce, Eric Baumer, Carl DiSalvo, and Phoebe Sengers. 2012. Sustainably unpersuaded: how persuasion narrows our vision of sustainability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 947–956.

[21] Charity Miles. 2021. Charity Miles. Apple App Store. https://apps.apple.com/us/app/charity-miles/id505253234

[22] Bagele Chilisa. 2019. *Indigenous research methodologies*. SAGE Publications, Thousand Oaks, CA.

[23] Janghee Cho, Laura Devendorf, and Stephen Voida. 2021. From The Art of Reflection to The Art of Noticing: A Shifting View of Self-Tracking Technologies' Role in Supporting Sustainable Food Practices. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 298, 7 pages. https://doi.org/10.1145/3411763.3451838

[24] Eun Kyoung Choe. 2014. *Designing self-monitoring technology to promote data capture and reflection*. Ph.D. Dissertation. University of Washington, Seattle, WA, USA.

[25] Eun Kyoung Choe, Saeed Abdullah, Mashfiqui Rabbi, Edison Thomaz, Daniel A Epstein, Felicia Cordeiro, Matthew Kay, Gregory D Abowd, Tanzeem Choudhury, James Fogarty, et al. 2017. Semi-automated tracking: a balanced approach for self-monitoring applications. *IEEE Pervasive Computing* 16, 1 (2017), 74–84.

[26] Eun Kyoung Choe, Bongshin Lee, and m.c. schraefel. 2015. Characterizing visualization insights from quantified selfers' personal data presentations. *IEEE computer graphics and applications* 35, 4 (2015), 28–37. https://doi.org/10.1109/MCG.2015.51

[27] 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*. ACM, New York, 173–182.

[28] Eun Kyoung Choe, Nicole B Lee, Bongshin Lee, Wanda Pratt, and Julie A Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 1143–1152.

[29] Chia-Fang Chung, Qiaosi Wang, Jessica Schroeder, Allison Cole, Jasmine Zia, James Fogarty, and Sean A. Munson. 2019. Identifying and Planning for Individualized Change: Patient-Provider Collaboration Using Lightweight Food Diaries in Healthy Eating and Irritable Bowel Syndrome. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 7 (March 2019), 27 pages. https://doi.org/10.1145/3314394

[30] Marc Clarà. 2015. What is reflection? Looking for clarity in an ambiguous notion. *Journal of teacher education* 66, 3 (2015), 261–271.

[31] Adrian K. Clear, Kirstie O'neill, Adrian Friday, and Mike Hazas. 2016. Bearing an Open "Pandora's Box": HCI for Reconciling Everyday Food and Sustainability. *ACM Trans. Comput.-Hum. Interact.* 23, 5, Article 28 (Oct. 2016), 25 pages. https://doi.org/10.1145/2970817

[32] Rob Comber, Jettie Hoonhout, Aart van Halteren, Paula Moynihan, and Patrick Olivier. 2013. *Food Practices as Situated Action: Exploring and Designing for Everyday Food Practices with Households*. Association for Computing Machinery, New York, NY, USA, 2457–2466. https://doi.org/10.1145/2470654.2481340

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

[34] Felicia Cordeiro, Elizabeth Bales, Erin Cherry, and James Fogarty. 2015. Rethinking the Mobile Food Journal: Exploring Opportunities for Lightweight Photo-Based Capture. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, 3207–3216. https://doi.org/10.1145/2702123.2702154

[35] Dan Cosley, Elizabeth Churchill, Jodi Forlizzi, and Sean A Munson. 2017. Introduction to this special issue on the lived experience of personal informatics. *Human–Computer Interaction* 32, 5-6 (2017), 197–207.

[36] Mayara Costa Figueiredo, Clara Caldeira, Tera L. Reynolds, Sean Victory, Kai Zheng, and Yunan Chen. 2017. Self-Tracking for Fertility Care: Collaborative Support for a Highly Personalized Problem. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 36 (Dec. 2017), 21 pages. https://doi.org/10.1145/3134671

[37] Mayara Costa Figueiredo, Thu Huynh, Anna Takei, Daniel A Epstein, and Yunan Chen. 2021. Goals, life events, and transitions: examining fertility apps for holistic health tracking. *JAMIA open* 4, 1 (2021), ooab013.

[38] Nediyana Daskalova, Eindra Kyi, Kevin Ouyang, Arthur Borem, Sally Chen, Sung Hyun Park, Nicole Nugent, and Jeff Huang. 2021. Self-E: Smartphone-Supported Guidance for Customizable Self-Experimentation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 227, 13 pages. https://doi.org/10.1145/3411764.3445100

[39] Nediyana Daskalova, Bongshin Lee, Jeff Huang, Chester Ni, and Jessica Lundin. 2018. Investigating the effectiveness of cohort-based sleep recommendations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 101.

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

[41] Anjali Devakumar, Jay Modh, Bahador Saket, Eric P. S. Baumer, and Munmun De Choudhury. 2021. A Review on Strategies for Data Collection, Reflection, and Communication in Eating Disorder Apps. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 547, 19 pages. https://doi.org/10.1145/3411764.3445670

[42] John Dewey. 1910. *How we think*. D. C. Health, Boston, MA. https://doi.org/10.1037/10903-000

[43] 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 (June 2021), 601–616. https://doi.org/10.1007/s00779-021-01573-w

[44] 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*. ACM, New York, 642–654.

[45] 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*. ACM, New York, 649–661.

[46] Chris Elsden, Abigail C. Durrant, and David S. Kirk. 2016. It's Just My History Isn't It? Understanding Smart Journaling Practices. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2819–2831. https://doi.org/10.1145/2858036.2858103

[47] 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.

[48] Chris Elsden, Mark Selby, Abigail Durrant, and David Kirk. 2016. Fitter, happier, more productive: what to ask of a data-driven life. *interactions* 23, 5 (2016), 45–45.

[49] 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

[50] 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*. ACM, New York, 1109–1113.

[51] Daniel A. Epstein, Jennifer H. Kang, Laura R. Pina, James Fogarty, and Sean A. Munson. 2016. Reconsidering the Device in the Drawer: Lapses as a Design

Opportunity in Personal Informatics. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. ACM, New York, 829–840. https://doi.org/10.1145/2971648.2971656

[52] Daniel A. Epstein, Nicole B. Lee, Jennifer H. Kang, Elena Agapie, Jessica Schroeder, Laura R. Pina, James Fogarty, Julie A. Kientz, and Sean Munson. 2017. Examining Menstrual Tracking to Inform the Design of Personal Informatics Tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, 6876–6888. https://doi.org/10.1145/3025453.3025635

[53] Daniel A Epstein, An Ping, James Fogarty, and Sean A Munson. 2015. A lived informatics model of personal informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, 731–742.

[54] FatSecret. 2021. Calorie Counter by FatSecret. Apple App Store. https://apps.apple.com/us/app/calorie-counter-by-fatsecret/id347184248

[55] Clayton Feustel, Shyamak Aggarwal, Bongshin Lee, and Lauren Wilcox. 2018. People Like Me: Designing for Reflection on Aggregate Cohort Data in Personal Informatics Systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 107.

[56] FITNESS22 LTD. 2021. Couch to 5K Runner. Apple App Store. https://apps.apple.com/gb/app/couch-to-5k-runner/id448357306

[57] FitNow. 2021. Lose It! — Calorie Counter. Apple App Store. https://apps.apple.com/us/app/lose-it-calorie-counter/id297368629

[58] Mary FitzGerald. 1994. Theories of reflection for learning. In *Reflective Practice in Nursing: The Growth of the Professional Practitioner*, Anthony Palmer, Sarah Burns, and Chris Bulman (Eds.). Blackwell Scientific Publications, Oxford, 63–84.

[59] 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*. ACM, New York, 216–223.

[60] FLO HEALTH, INC. 2021. Flo Period & Ovulation Tracker. Apple App Store. https://apps.apple.com/us/app/flo-period-ovulation-tracker/id1038369065

[61] Forward Human, Inc. 2021. Jour: Daily Self-Care Journal. Apple App Store. https://apps.apple.com/us/app/jour-daily-self-care-journal/id1439590239

[62] Elliot G. Mitchell, Elizabeth M. Heitkemper, Marissa Burgermaster, Matthew E. Levine, Yishen Miao, Maria L. Hwang, Pooja M. Desai, Andrea Cassells, Jonathan N. Tobin, Esteban G. Tabak, David J. Albers, Arlene M. Smaldone, and Lena Mamykina. 2021. From Reflection to Action: Combining Machine Learning with Expert Knowledge for Nutrition Goal Recommendations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 206, 17 pages. https://doi.org/10.1145/3411764.3445555

[63] Eva Ganglbauer, Geraldine Fitzpatrick, and Florian Güldenpfennig. 2015. Why and what did we throw out? Probing on reflection through the food waste diary. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, NEw York, 1105–1114.

[64] Maliheh Ghajargar. 2018. *Designing Tools for Reflection*. Ph.D. Dissertation. Politecnico di Torino. https://doi.org/10.6092/polito/porto/2702550

[65] Glow. 2021. Nurture: Pregnancy + Baby App. Apple App Store. https://apps.apple.com/us/app/nurture-pregnancy-baby-app/id882398397#?platform=iphone

[66] Glow. 2021. Period Tracker - Eve. Apple App Store. https://apps.apple.com/us/app/period-tracker-eve/id1002275138#?platform=iphone

[67] Mitchell Gordon, Tim Althoff, and Jure Leskovec. 2019. Goal-Setting And Achievement In Activity Tracking Apps: A Case Study Of MyFitnessPal. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. Association for Computing Machinery, New York, NY, USA, 571–582. https://doi.org/10.1145/3308558.3313432

[68] 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*. ACM, New York, 1305–1316.

[69] Lisa Graham, Anthony Tang, and Carman Neustaedter. 2016. Help Me Help You: Shared Reflection for Personal Data. In *Proceedings of the 19th International Conference on Supporting Group Work* (Sanibel Island, Florida, USA) *(GROUP '16)*. Association for Computing Machinery, New York, NY, USA, 99–109. https://doi.org/10.1145/2957276.2957293

[70] Lars Grammel, Melanie Tory, and Margaret-Anne Storey. 2010. How information visualization novices construct visualizations. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 943–952.

[71] Andrea Grimes and Richard Harper. 2008. Celebratory Technology: New Directions for Food Research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 467–476. https://doi.org/10.1145/1357054.1357130

[72] Shad Gross, Jeffrey Bardzell, Shaowen Bardzell, and Michael Stallings. 2017. Persuasive anxiety: designing and deploying material and formal explorations of personal tracking devices. *Human–Computer Interaction* 32, 5-6 (2017), 297–334.

[73] Rebecca Gulotta, Jodi Forlizzi, Rayoung Yang, and Mark Wah Newman. 2016. Fostering engagement with personal informatics systems. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, New York, 286–300.

[74] Jürgen Habermas. 1987. *Knowledge and human interests*. Polity Press, Cambridge, UK.

[75] Lars Hallnäs and Johan Redström. 2001. Slow technology–designing for reflection. *Personal and ubiquitous computing* 5, 3 (2001), 201–212.

[76] Steve Harrison, Phoebe Sengers, and Deborah Tatar. 2011. Making epistemological trouble: Third-paradigm HCI as successor science. *Interacting with Computers* 23, 5 (2011), 385–392.

[77] Giyoo Hatano and Kayoko Inagaki. 1984. Two courses of expertise. *RESEARCH AND CLINICAL CENTER FOR CHILD DEVELOPMENT Annual Report* 6 (1984), 27–36.

[78] Neville Hatton and David Smith. 1995. Reflection in teacher education: Towards definition and implementation. *Teaching and teacher education* 11, 1 (1995), 33–49.

[79] Dize Hilviu and Amon Rapp. 2015. Narrating the quantified self. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, New York, 1051–1056.

[80] Huckleberry Labs Inc. 2021. Huckleberry: Baby & Child. Apple App Store. https://apps.apple.com/us/app/huckleberry-baby-child/id1169136078

[81] Ravi Karkar, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, Roger Vilardaga, and Jasmine Zia. 2017. *TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers*. Association for Computing Machinery, New York, NY, USA, 6850–6863. https://doi.org/10.1145/3025453.3025480

[82] Maria Karyda, Elisa D Mekler, and Andrés Lucero. 2021. Data Agents: Promoting Reflection through Meaningful Representations of Personal Data in Everyday Life. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 367, 11 pages. https://doi.org/10.1145/3411764.3445112

[83] Stephen Kemmis. 1985. Action research and the politics of reflection. In *Reflection: Turning experience into learning*, David Boud, Rosemary Keogh, and David Walker (Eds.). RoutledgeFarmer, London, 139–163.

[84] 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.

[85] Rohit Ashok Khot, Larissa Hjorth, and Florian 'Floyd' Mueller. 2014. Understanding Physical Activity through 3D Printed Material Artifacts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3835–3844. https://doi.org/10.1145/2556288.2557144

[86] Vera Khovanskaya, Eric PS Baumer, Dan Cosley, Stephen Voida, and Geri Gay. 2013. Everybody knows what you're doing: a critical design approach to personal informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 3403–3412.

[87] Nam Wook Kim, Hyejin Im, Nathalie Henry Riche, Alicia Wang, Krzysztof Gajos, and Hanspeter Pfister. 2019. DataSelfie: Empowering People to Design Personalized Visuals to Represent Their Data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 79.

[88] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 67.

[89] Kindara, Inc. 2021. Kindara: Fertility Tracker. Apple App Store. https://apps.apple.com/us/app/kindara-fertility-tracker/id522674372#?platform=iphone

[90] 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 Conference on Designing Interactive Systems*. ACM, New York, 881–894.

[91] Rafal Kocielnik, Gary Hsieh, and Daniel Avrahami. 2018. Helping Users Reflect on Their Own Health-Related Behaviors. In *Studies in Conversational UX Design*, Robert J. Moore, Margaret H. Szymanski, Raphael Arar, and Guang-Jie Ren (Eds.). Springer International, Cham, Switzerland, 85–115.

[92] 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), 70.

[93] Lenneke Kuijer and Elisa Giaccardi. 2018. *Co-Performance: Conceptualizing the Role of Artificial Agency in the Design of Everyday Life*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/

3173574.3173699
[94] Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H Nguyen. 2015. Why we use and abandon smart devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, 635–646.
[95] Min Kyung Lee, Junsung Kim, Jodi Forlizzi, and Sara Kiesler. 2015. Personalization revisited: a reflective approach helps people better personalize health services and motivates them to increase physical activity. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, New York, 743–754.
[96] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, 557–566.
[97] Ian Li, Anind K Dey, and Jodi Forlizzi. 2012. Using context to reveal factors that affect physical activity. *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 1 (2012), 7.
[98] Zilu Liang, Bernd Ploderer, Wanyu Liu, Yukiko Nagata, James Bailey, Lars Kulik, and Yuxuan Li. 2016. SleepExplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors. *Personal and Ubiquitous Computing* 20, 6 (2016), 985–1000.
[99] Lifesum AB. 2021. Lifesum: Healthy Eating. Apple App Store. https://apps.apple.com/us/app/lifesum-diet-macro-tracker/id286906691
[100] Catherine Y Lim, Andrew BL 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*. ACM, New York, 655.
[101] Xiaodong Lin, Cindy Hmelo, Charles K Kinzer, and Teresa J Secules. 1999. Designing technology to support reflection. *Educational Technology Research and Development* 47, 3 (1999), 43–62.
[102] Silvia Lindtner, Shaowen Bardzell, and Jeffrey Bardzell. 2018. Design and Intervention in the Age of No Alternative. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 109.
[103] Szu-Yu Cyn Liu, Jen Liu, Kristin Dew, Patrycja Zdziarska, Maya Livio, and Shaowen Bardzell. 2019. Exploring Noticing as Method in Design Research. In *Companion Publication of the 2019 on Designing Interactive Systems*. ACM, New York, 377–380.
[104] Edwin A Locke and Gary P Latham. 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist* 57, 9 (2002), 705.
[105] Edwin A Locke and Gary P Latham. 2006. New directions in goal-setting theory. *Current directions in psychological science* 15, 5 (2006), 265–268.
[106] Yuhan Luo, Bongshin Lee, and Eun Kyoung Choe. 2020. TandemTrack: Shaping Consistent Exercise Experience by Complementing a Mobile App with a Smart Speaker. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. ACM, New York, 1–13. https://doi.org/10.1145/3313831.3376616
[107] Yuhan Luo, Peiyi Liu, and Eun Kyoung Choe. 2019. Co-Designing Food Trackers with Dietitians: Identifying Design Opportunities for Food Tracker Customization. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM Press, New York, 592.
[108] Deborah Lupton. 2014. Apps as artefacts: Towards a critical perspective on mobile health and medical apps. *Societies* 4, 4 (2014), 606–622.
[109] Deborah Lupton. 2014. Self-tracking modes: Reflexive self-monitoring and data practices. SSRN. , 19 pages. https://doi.org/10.2139/ssrn.2483549
[110] Deborah Lupton. 2019. The thing-power of the human-app health assemblage: thinking with vital materialism. *Social Theory & Health* 17, 2 (2019), 125–139.
[111] 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
[112] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-Rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 72 (nov 2019), 23 pages. https://doi.org/10.1145/3359174
[113] Jack Mezirow. 1981. A critical theory of adult learning and education. *Adult education* 32, 1 (1981), 3–24.
[114] Jack Mezirow. 1991. *Transformative dimensions of adult learning*. Jossey-Bass, San Francisco.
[115] Matthew B Miles, A Michael Huberman, Michael A Huberman, and Michael Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. SAGE Publications, Thousand Oaks, CA.
[116] Mint.com. 2021. Mint: Budget & Expense Tracker. Apple App Store. https://apps.apple.com/us/app/mint-personal-finance-money/id300238550
[117] Ine Mols, Elise van den Hoven, and Berry Eggen. 2016. Informing Design for Reflection: An Overview of Current Everyday Practices. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (Gothenburg, Sweden) *(NordiCHI '16)*. Association for Computing Machinery, New York, NY, USA, Article 21, 10 pages. https://doi.org/10.1145/2971485.2971494
[118] Ine Mols, Elise Van Den Hoven, and Berry Eggen. 2016. Technologies for everyday life reflection: Illustrating a design space. In *Proceedings of the Tenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, New York, 53–61.
[119] Sean A Munson. 2018. Rethinking assumptions in the design of health and wellness tracking tools. *interactions* 25, 1 (2018), 62–65.
[120] Sean A Munson and Sunny Consolvo. 2012. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*. ICST, Gent, Belgium, 25–32. https://doi.org/10.4108/icst.pervasivehealth.2012.248691
[121] Elizabeth L. Murnane, Xin Jiang, Anna Kong, Michelle Park, Weili Shi, Connor Soohoo, Luke Vink, Iris Xia, Xin Yu, John Yang-Sammataro, Grace Young, Jenny Zhi, Paula Moya, and James A. Landay. 2020. Designing Ambient Narrative-Based Interfaces to Reflect and Motivate Physical Activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. ACM, New York, 1–14. https://doi.org/10.1145/3313831.3376478
[122] Elizabeth L Murnane, Tara G Walker, Beck Tench, Stephen Voida, and Jaime Snyder. 2018. Personal informatics in interpersonal contexts: towards the design of technology that supports the social ecologies of long-term mental health management. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 127.
[123] MyFitnessPal, Inc. 2021. MyFitnessPal. Apple App Store. https://apps.apple.com/us/app/myfitnesspal/id341232718
[124] MyNetDiary, Inc. 2021. Calorie Counter - MyNetDiary. Apple App Store. https://apps.apple.com/us/app/calorie-counter-mynetdiary/id287529757
[125] Jasmin Niess and Paweł W. Woundefinedniak. 2018. *Supporting Meaningful Personal Fitness: The Tracker Goal Evolution Model*. ACM, New York, 1–12. https://doi.org/10.1145/3173574.3173745
[126] Nighp Software LLC. 2021. Baby Tracker - Newborn Log. Apple App Store. https://apps.apple.com/us/app/baby-tracker-newborn-log/id779656557#?platform=iphone
[127] Helen Noble and Joanna Smith. 2015. Issues of validity and reliability in qualitative research. *Evidence-based nursing* 18, 2 (2015), 34–35.
[128] Noom, Inc. 2021. Noom. Apple App Store. https://apps.apple.com/us/app/noom/id634598719
[129] Don Norman. 2014. *Things that make us smart: Defending human attributes in the age of the machine*. Diversion Books, New York.
[130] Dietmar Offenhuber. 2019. Data by proxy—Material traces as autographic visualizations. *IEEE transactions on visualization and computer graphics* 26, 1 (2019), 98–108.
[131] Ovulation Calculator Pty Ltd. 2017. Ovulation Calculator Fertile Tracker & Calendar OC. Apple App Store. https://apps.apple.com/us/app/ovulation-calculator-fertile-tracker-calendar-oc/id1010506484
[132] Ovuline, Inc. 2021. Ovia Fertility & Cycle Tracker. Apple App Store. https://apps.apple.com/us/app/ovia-fertility-cycle-tracker/id570244389
[133] Ovuline, Inc. 2021. Ovia Parenting & Baby Tracker. Apple App Store. https://apps.apple.com/us/app/ovia-parenting-baby-tracker/id1106614359
[134] Pacer, Inc. 2021. Pacer Pedometer & Step Tracker. Apple App Store. https://apps.apple.com/us/app/pacer-pedometer-step-tracker/id600446812
[135] Andrea Grimes Parker. 2014. Reflection-through-Performance: Personal Implications of Documenting Health Behaviors for the Collective. *Personal Ubiquitous Comput.* 18, 7 (Oct. 2014), 1737–1752. https://doi.org/10.1007/s00779-014-0780-5
[136] Phoenix Games LLP. 2018. Drink Water Reminder N Tracker. Apple App Store. https://apps.apple.com/us/app/drink-water-reminder-n-tracker/id870372885#?platform=iphone
[137] Bernd Ploderer, Wolfgang Reitberger, Harri Oinas-Kukkonen, and Julia van Gemert-Pijnen. 2014. Social interaction and reflection for behaviour change.
[138] Michael Prilla, Martin Degeling, and Thomas Herrmann. 2012. Collaborative Reflection at Work: Supporting Informal Learning at a Healthcare Workplace. In *Proceedings of the 17th ACM International Conference on Supporting Group Work* (Sanibel Island, Florida, USA) *(GROUP '12)*. Association for Computing Machinery, New York, NY, USA, 55–64. https://doi.org/10.1145/2389176.2389185
[139] Nora Ptakauskaite, Anna L. Cox, and Nadia Berthouze. 2018. Knowing What You're Doing or Knowing What to Do: How Stress Management Apps Support Reflection and Behaviour Change. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188648
[140] Amon Rapp. 2018. *Gamification for Self-Tracking: From World of Warcraft to the Design of Personal Informatics Systems*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3173574.3173654
[141] 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.

[142] Amon Rapp, Maurizio Tirassa, and Lia Tirabeni. 2019. Rethinking Technologies for Behavior Change: A View from the Inside of Human Change. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 4 (2019), 22.

[143] Recovery Record. 2021. RR: Eating Disorder Management. Apple App Store. https://apps.apple.com/us/app/rr-eating-disorder-management/id457360959

[144] Saeyoung Rho, Injung Lee, Hankyung Kim, Jonghyuk Jung, Hyungi Kim, Bong Gwan Jun, and Youn-kyung Lim. 2017. FutureSelf: What Happens When We Forecast Self-Trackers? Future Health Statuses?. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 637–648. https://doi.org/10.1145/3064663.3064676

[145] Verónica Rivera-Pelayo, Angela Fessl, Lars Müller, and Viktoria Pammer. 2017. Introducing Mood Self-Tracking at Work: Empirical Insights from Call Centers. *ACM Trans. Comput.-Hum. Interact.* 24, 1, Article 3 (Feb. 2017), 28 pages. https://doi.org/10.1145/3014058

[146] 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*. ACM, New York, 1163–1172.

[147] Maria Roussou, Sara Perry, Akrivi Katifori, Stavros Vassos, Angeliki Tzouganatou, and Sierra McKinney. 2019. Transformation through Provocation?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. ACM, New York, 1–13. https://doi.org/10.1145/3290605.3300857

[148] 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)*. ACM, New York, 1–13. https://doi.org/10.1145/3173574.3173883

[149] Sanvello Health Inc. 2021. Sanvello: Anxiety & Depression. Apple App Store. https://apps.apple.com/us/app/sanvello-anxiety-depression/id922968861#?platform=iphone

[150] Donald A Schön. 1987. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. 1987.* Jossey Bass, San Francisco.

[151] Donald A Schön. 2017. *The reflective practitioner: How professionals think in action.* Routledge, New York.

[152] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. 2005. Reflective design. In *Proceedings of the 4th Decennial Conference on Critical Computing*. ACM, New York, 49–58.

[153] Skimble. 2021. Workout Trainer: fitness coach. Apple App Store. https://apps.apple.com/us/app/workout-trainer-fitness-coach/id395686735

[154] Sleep Cycle AB. 2021. Sleep Cycle - Sleep Tracker. Apple App Store. https://apps.apple.com/us/app/sleep-cycle-sleep-tracker/id320606217

[155] Petr Slovak, Christopher Frauenberger, and Geraldine Fitzpatrick. 2017. Reflective practicum: A framework of sensitising concepts to design for transformative reflection. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, New York, 2696–2707.

[156] SnoreLab Ltd. 2021. SnoreLab : Record Your Snoring. Apple App Store. https://apps.apple.com/us/app/snorelab-record-your-snoring/id529443604

[157] Katta Spiel, Fares Kayali, Louise Horvath, Michael Penkler, Sabine Harrer, Miguel Sicart, and Jessica Hammer. 2018. Fitter, Happier, More Productive? The Normative Ontology of Fitness Trackers. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. ACM, New York, 1–10. https://doi.org/10.1145/3170427.3188401

[158] StepsApp GMBH. 2021. StepsApp Pedometer. Apple App Store. https://apps.apple.com/us/app/stepsapp-pedometer/id1037595083

[159] Strava, Inc. 2021. Strava: Run, Ride, Swim. Apple App Store. https://apps.apple.com/us/app/strava-run-ride-swim/id426826309

[160] S Shyam Sundar. 2008. Self as source: Agency and customization in interactive media. In *Mediated Interpersonal Communication*, Elly A Konijn, Sonja Utz, Martin Tanis, and Susan B Barnes (Eds.). Routledge, New York, 58–74. https://doi.org/10.4324/9780203926864

[161] Melanie Swan. 2013. The quantified self: Fundamental disruption in big data science and biological discovery. *Big data* 1, 2 (2013), 85–99.

[162] Tamtris Web Services Inc. 2021. Fertility Friend FF App. Apple App Store. https://apps.apple.com/us/app/fertility-friend-ff-app/id443919067

[163] Charlene Tan. 2020. Revisiting Donald Schön's notion of reflective practice: a Daoist interpretation. *Reflective Practice* 21, 5 (2020), 686–698.

[164] The Lampo Group Incorporated. 2021. EveryDollar: Budget Your Money. Apple App Store. https://apps.apple.com/us/app/everydollar-budget-your-money/id942571931#?platform=iphone

[165] Jakob Tholander and Stina Nylander. 2015. *Snot, Sweat, Pain, Mud, and Snow: Performance and Experience in the Use of Sports Watches.* Association for Computing Machinery, New York, NY, USA, 2913–2922. https://doi.org/10.1145/2702123.2702482

[166] Neil Thompson and Jan Pascal. 2012. Developing critically reflective practice. *Reflective practice* 13, 2 (2012), 311–325.

[167] Alice Thudt, Uta Hinrichs, Samuel Huron, and Sheelagh Carpendale. 2018. Self-reflection and personal physicalization construction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM Press, New York, 154.

[168] Martin Tisné. 2019. Data isn't the new oil, it's the new CO2. https://luminategroup.com/posts/blog/data-isnt-the-new-oil-its-the-new-co2

[169] Truebill, Inc. 2021. Truebill Budget & Bill Tracker. Apple App Store. https://apps.apple.com/us/app/truebill-budget-bill-tracker/id1130616675

[170] Anna Lowenhaupt Tsing. 2015. *The mushroom at the end of the world: On the possibility of life in capitalist ruins.* Princeton University Press, Princeton, NJ.

[171] Under Armour, Inc. 2021. Map My Fitness by Under Armour. Apple App Store. https://apps.apple.com/us/app/map-my-fitness-by-under-armour/id298903147

[172] Under Armour, Inc. 2021. Map My Run by Under Armour. Apple App Store. https://apps.apple.com/us/app/map-my-run-by-under-armour/id291890420

[173] Elisabeth Kersten van Dijk, Wijnand IJsselsteijn, and Joyce Westerink. 2016. Deceptive Visualizations and User Bias: A Case for Personalization and Ambiguity in PI Visualizations. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 588–593. https://doi.org/10.1145/2968219.2968326

[174] Viktor Sharov. 2021. My Water: Daily Drink Tracker. Apple App Store. https://apps.apple.com/us/app/my-water-daily-drink-tracker/id964748094#?platform=iphone

[175] Sue White, Jan Fook, and Fiona Gardner. 2006. *Critical reflection in health and social care.* Open University Press, Maidenhead, UK.

[176] Winkpass Creations, Inc. 2021. iPeriod Lite Period Tracker. Apple App Store. https://apps.apple.com/us/app/iperiod-lite-period-tracker/id340757216

[177] Gary Wolf. 2009. Know Thyself: Tracking Every Facet of Life, from Sleep to Mood to Pain, 24/7/365. *WIRED* 17.07 (June 2009), 92–95. http://www.wired.com/medtech/health/magazine/17-07/lbnp_knowthyself?currentPage=all

[178] Wombat Apps LLC. 2021. Carb Manager—Keto Diet Tracker. Apple App Store. https://apps.apple.com/us/app/carb-manager-keto-diet-app/id410089731#?platform=iphone

[179] Paweł W. Woundefinedniak, Przemysław Piotr Kucharski, Maartje M.A. de Graaf, and Jasmin Niess. 2020. Exploring Understandable Algorithms to Suggest Fitness Tracker Goals That Foster Commitment. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 35, 12 pages. https://doi.org/10.1145/3419249.3420131

[180] Paweł W. Woundefinedniak, Przemysław Piotr Kucharski, Maartje M.A. de Graaf, and Jasmin Niess. 2020. Exploring Understandable Algorithms to Suggest Fitness Tracker Goals That Foster Commitment. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. ACM, New York, Article 35, 12 pages. https://doi.org/10.1145/3419249.3420131

[181] Renwen Zhang, Kathryn E. Ringland, Melina Paan, David C. Mohr, and Madhu Reddy. 2021. Designing for Emotional Well-Being: Integrating Persuasion and Customization into Mental Health Technologies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 542, 13 pages. https://doi.org/10.1145/3411764.3445771