Understanding and Supporting Self-Tracking App Selection

JONG HO LEE, University of Maryland, College Park, USA
JESSICA SCHROEDER, University of Washington, USA
DANIEL A. EPSTEIN, University of California, Irvine, USA

People often face barriers to selecting self-tracking tools that support their goals and needs, resulting in tools not meeting their expectations and ultimately abandonment. We therefore examine how people approach selecting self-tracking apps and investigate how technology can better support the process. Drawing on past literature on how people select and perceive the features of commercial and research tracking tools, we surface seven attributes people consider during selection, and design a low-fidelity prototype of an app store that highlights these attributes. We then conduct semi-structured interviews with 18 participants to further investigate what people consider during selection, how people select self-tracking apps, and how surfacing tracking-related attributes could better support selection. We find that people often prioritize features related to self-tracking during selection, such as approaches to collecting and reflecting on data, and trial apps to determine whether they would suit their needs. Our results also show potential for technology surfacing how apps support tracking to reduce barriers to selection. We discuss future opportunities for improving self-tracking app selection, such as ways to enhance existing self-tracking app distribution platforms to enable people to filter and search apps by desirable features.

CCS Concepts: • **Human-centered computing → Empirical studies in ubiquitous and mobile computing**.

Additional Key Words and Phrases: Personal Informatics; Quantified Self; Self-Tracking; Selection; App Stores

ACM Reference Format:
Jong Ho Lee, Jessica Schroeder, and Daniel A. Epstein. 2021. Understanding and Supporting Self-Tracking App Selection. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 166 (December 2021), 25 pages. https://doi.org/10.1145/3494980

1 INTRODUCTION

The advent of smartphones introduced apps which allow people to track aspects of their personal lives. Many U.S. adults have used a self-tracking app to track their physical activity (44%), diet (42%), sleep (25%), or mental health (23%) [68], with the number of people using health and fitness apps growing from 62.7 million in 2018 to 86.3 million in 2020 [69]. Self-tracking apps have the potential to improve people's lives in many aspects, from getting better sleep [20, 21] to managing chronic symptoms [64, 65] and achieving personal fitness goals [58].

However, substantial research has shown that people often abandon these tracking technologies, or personal informatics systems, prior to achieving the benefits they desire because the tools they select do not align with their expectations [14, 25, 28, 47]. For example, some people find that tracking requires more effort than they anticipated [25, 47], find the tool they selected does not support tracking the data they desire [14, 28, 47, 67], or find the tools they selected a bad aesthetic or stylistic fit [27, 36, 47]. Unfortunately, people also face barriers when selecting self-tracking tools. A common barrier is the sheer number of consumer tools available. Just among apps, over 318,000 health apps were available in 2017, with 200 more being added each day [40], making it

Authors' addresses: Jong Ho Lee, University of Maryland, College Park, College Park, USA, jlee29@umd.edu; Jessica Schroeder, University of Washington, Seattle, USA, jesscs@cs.washington.edu; Daniel A. Epstein, University of California, Irvine, Irvine, USA, epstein@ics.uci.edu.

2474-9567/2021/12-ART166 $15.00
https://doi.org/10.1145/3494980

challenging to differentiate between and choose from available options. Another barrier is people's diverse needs for self-tracking apps, which tend to influence the tools they choose to track with [28, 65]. People have various tracking styles and goals, including behavior change, curiosity, and awareness [62], and people's goals evolve as they learn more about their habits and what technology can support tracking [3, 28, 46]. The barriers people face when selecting self-tracking tools can be particularly detrimental to their overall self-tracking process because they often cascade. Problems encountered early on during self-tracking, including selecting an appropriate tool, hinder later efforts to derive value and self-understanding from self-tracked data [50].

Although previous work in mobile app selection has identified general factors influencing adoption such as an app's visual appearance [22, 41, 54], the entertainment value of an app [9, 35, 54], or general user experience [1, 10, 22, 54], whether and to what extent these findings explain people's decision-making when selecting self-tracking apps is unclear. For example, after use, people consider the presence or implementation of specific features, such as whether collecting certain data is supported, alignment with self-tracking goals, or how apps support reflection [14, 25, 28, 47]. Although research has suggested that people often assess whether a tool will address their needs prior to deciding to use it [28, 50], we have relatively little understanding of what criteria people use to make that assessment, and how those criteria align with or differ from criteria used to select other mobile apps. Additionally, the personal informatics literature has offered relatively little guidance on how technology can support people's selection practices [24]. Developing this understanding can help researchers and practitioners offer advice or create tools to support people in selecting tracking tools which are appropriate for their needs.

To understand how people approach selecting tracking tools and how technology can better support them in assessing tools, we first reviewed prior personal informatics literature to understand what people typically consider and prioritize in self-tracking tools. We identified seven *attributes* that people consider, including what data is collected and approaches to presenting feedback. We then redesigned an app store in a low-fidelity prototype to expose the tracking-related attributes and enable searching by those attributes. Finally, we conducted semi-structured interviews with 18 participants to better understand how people search and select self-tracking apps, using the prototype to elicit feedback on the utility of the attributes for supporting selection. We contribute:

- An understanding of what people consider when selecting personal tracking tools. Prior literature on people's adoption and use of commercial and research tools reveal seven *attributes* of self-tracking apps that people consider: *data collected*, *feedback provided*, *goal-setting capabilities*, *privacy*, *social opportunities*, *style*, and *convenience*. Our interviews confirm that people frequently considered these attributes, alongside factors broadly applicable to apps such as an app's reputation, developer support, and cost.
- An understanding of how people approach selecting personal tracking tools in practice. When gathering information about self-tracking apps, people frequently *trial*, or experiment with, many apps prior to settling on one or couple to use for self-understanding. When trialing, some participants collected data in ways similar to how they would track as if they would use the app long-term, while others recorded inconsequential data to get firsthand experience at what tracking might be like were they to select that app.
- Design considerations for how technology can better support people in selecting personal tracking tools. Participant feedback on exposing the attributes of self-tracking apps through searches, labels, and filters suggest that technology surfacing how tracking tools practically support tracking can potentially help reduce barriers to selection. When implementing such ideas, however, technology should consider challenges in interpreting tags and potentially subjecting people to more information about apps than they desire.

2 BACKGROUND AND RELATED WORK

Our understanding and support of self-tracking selection builds on past research discussing selection practices in commodity apps, personal informatics preparation and selection, and technology approaches to supporting those practices. Here we provide context in how prior research examined mobile app selection and our current selection approaches in personal informatics. Substantial work also examines people's experiences selecting commercial tracking tools and reflecting on their choices, which we further discuss in Section 3.

2.1 Selecting Smartphone and Self-Tracking Apps

People's choice of a self-tracking tool often impacts their ability to reach their goals of gaining self-understanding or self-knowledge. Li et al. characterize the first stage of tracking as *preparation* [50], which Epstein et al. later suggest can be considered as separate actions of *deciding* to track and *selecting* a tool [28]. Li et al. describe how barriers faced throughout the process of self-tracking cascade, with barriers in earlier stages having consequences for people's later efforts to understand and act on their data [50]. Reinforcing the importance of this cascading effect, research examining why people abandon self-tracking technology identify mismatches between what tools support and people's expectations for them [14, 25, 47]. After using a tracking tool for some time, people often find that the tools they select do not support collection of the data they are most interested in [14, 47] or find that they desire more granular or accurate data than tools are able to provide [17, 25, 67]. Others find the method of collecting data more burdensome than they desired [25, 47]. People sometimes indicate wanting specific features that the tool they selected does not support, such as dedicated social feeds [67] or disliking traits which apps do include, such as sharing data with third parties [25].

Prior work has characterized that people often go through three phases when adopting mobile apps (regardless of domain): Search, Decide, and Try [1]. When *Search*ing, people consider their needs for a suitable app and what sources of information would be helpful for learning about apps (e.g., app stores, social media). After searching, people *Decide* whether to install an app or multiple apps. People then typically Try or *Trial* the app or apps for a short amount of time, choosing whether to adopt the app or reject the app based on the trial. People's process of trying apps is relatively similar to their use of trials of other products, evaluating the experiential (e.g., how effective the product was) and non-experiential (e.g., what emotions use of the product produced) aspects of the product [44], often trialing multiple apps at a time [49, 60, 72]. When people reject apps, they often return to searching for an app, having built more of an understanding of their goals for an app [22, 39]. The decision to install an app does *not* imply adoption; many people end up rejecting the app or search for a different one after the trial period [1, 38, 66]. Although prior related research contributed to the understanding of how people select apps in general, relatively little is known about how the unique characteristics of self-tracking (e.g., the need to collect and reflect on data) influence how people undertake searching for, deciding on, and trialing tracking tools specifically. Understanding how people interpret and evaluate what data an app collects and how they might reflect on that data can point to useful directions for technology to better assist in self-tracking tool selection.

When searching for apps, people gather information from various sources. People often learn about apps via recommended lists on app stores [41, 70] and evaluate apps' reputation from star ratings and reviews [1, 22, 35]. Other sources of information include app websites [70], close friends or family members [9, 35, 41], looser connections on social media [70], experts like physicians [70], printed periodicals and books [70], television [70], or recommendations from within other apps [49]. In self-tracking, people's motivations for tracking can impact the how they select tools; for example, people interested in tracking for self-improvement are more likely to carefully research potential tools and consider their features, while people with curiosity or record-keeping goals are more likely select tools out of convenience (e.g., what is installed by default), social influence (e.g., what friends or family use), or requirement (e.g., what provides credit in a workplace wellness program) [28].

Prior literature has characterized factors and aspects people consider when deciding whether to adopt an app. Research has primarily discussed overarching principles people consider across apps, including substantial work highlighting that the overall user experience of an app impacts interest and adoption [1, 10, 54, 70]. People similarly find the visual appeal of an app important, such as app icons and names [22, 39, 66] or of the app itself [9, 41, 66]. People also often consider privacy, with researchers finding that people being aware of apps' privacy policies influenced adoption [1, 22, 59, 66, 70]. Other notable factors include how entertaining the app is [9, 35, 66], how well it aligns with an individual's personality [59, 73], and the developer's reputation [22, 41, 66, 70].

Research across app domains indicates that people consider the content and features of apps, and how that content aligns with their needs. For example, some people select health apps based on whether they include features like workout routines [41], ability to self-diagnose based on symptoms [70], or the ability to connect apps with social media accounts [9]. Research has highlighted that people consider whether the health recommendations in apps are informed or validated by domain experts prior to adoption [66, 70]. The importance of presence or implementation of features when evaluating apps are typically domain-specific. For example, many of the feature considerations people make when selecting apps for a specific purpose are often less relevant to other kinds of apps and circumstances, such as a need to support different accounts when selecting password managers [1]. Understanding what aspects and features people consider and prioritize when selecting self-tracking apps deepens our understanding of what value and knowledge people aim to derive. We therefore investigate:

- RQ1) What about personal tracking apps do people consider and prioritize during selection?
- RQ2) What processes and procedures do people undertake when selecting personal tracking apps?

2.2 Addressing Barriers to App and Self-Tracking Selection

To address barriers to app selection of apps generally, research has proposed improvements to app stores, a frequent resource for searching for and learning about apps. For example, research has suggested making app recommendations, leveraging the preferences and experiences of other mobile app users [51, 75]. Research has also suggested allowing people to search apps by features, in particular by surfacing privacy concerns [30, 43]. However, the general nature of these strategies offers only high-level guidance of how app stores can support selection in personal informatics, with little focus on how app stores could surface the self-tracking features which apps include and how they support those features.

As an alternative to supporting selection among multiple self-tracking tools, personal informatics research has explored the alternative of increasing flexibility in tracking tools themselves, providing freedom to customize data collection within a single tool. Systems have aimed to support this flexibility, like OmniTrack allowing people to create their own trackers by customizing the frequency and fidelity of tracking items [46] and Trackly allowing configuration of parameters they tracked and how they are visualized [3]. However, research has suggested that commercial self-tracking apps and technology are often too rigid to support people's diverse and changing needs [4, 64] and people find that tools do not currently support the levels of flexibility they desire [11, 36, 46]. Research has argued that a single app is unlikely to support the varied preferences and goals that people have as they track, instead suggesting a need to leverage ecosystems of tools [55, 66]. To address this gap in understanding how technology can promote selection among multiple self-tracking apps, we investigate:

- RQ3) How can technology better support people in selecting personal tracking apps?

3 CHARACTERIZING IMPORTANT ATTRIBUTES OF SELF-TRACKING APPS FOR SELECTION

To gain insight on what attributes of self-tracking apps people consider when selecting (RQ1), we first reviewed past personal informatics literature on people's experiences with commercial self-tracking systems and research prototypes, and their motivations for adopting them. Based on this review, we identified attributes of self-tracking features and factors that may help people understand whether a self-tracking tool serves their needs.

3.1 Method

We first looked at related publications based on our previous experience studying the personal informatics literature. We specifically sought out publications which discussed: 1) why people stop or continue using self-tracking tools; 2) people's perspectives on the design of the self-tracking tools they were using; 3) how people

Table 1. Our examination of prior literature on self-tracking identified seven attributes of self-tracking apps people consider or might want to consider when selecting a tracking tool.

Attributes	Factors	Examples	Described In
Data Collected	What it collects	Run mileage, mood, screen time	[6, 12, 14, 15, 20, 28, 36, 46, 47, 53, 65, 67]
	How it is collected	Manual self-report, automatically sensed	[12, 17, 26, 46, 47, 50, 53, 74]
Feedback Provided	Kind of feedback	Aggregate daily totals, recommendations	[6, 14, 20, 26, 42, 46, 47, 53, 65]
	How it is presented	Graphs, numeric summary, text summary	[6, 12, 20, 21, 26, 27, 42, 46, 47, 53, 65]
Goal-setting Capabilities	Kind of goal	Time goal, distance goal, no goal support	[12, 15, 20, 21, 26, 31, 32, 34, 42, 53, 58, 62, 65]
	Goal duration	Daily, weekly, custom duration	[15, 19, 23, 31, 32, 34, 46, 58]
	Goal configurability	Can be edited, not configurable	[15, 19, 23, 31, 46, 48, 56, 65]
Privacy	Data disclosure	Shares data with developer, does not share	[13, 25, 31, 45]
	History	Can clear history, cannot clear history	[25, 45]
Social Opportunities	Social features	Following friends, in-app chat, share goals	[14, 15, 26, 31, 33, 52, 56, 62, 67]
	What is shared	Goals, data, rewards	[15, 26, 31, 36, 56, 62]
Style	Aesthetics and feel	Sporty, serious, playful	[27, 36, 47, 67]
Convenience	Third-party support	Syncs with Apple Health, can export data	[6, 12, 14, 26, 28, 47, 50]
	Notifications	Customizable notifications	[5, 6, 21, 25, 26, 42, 46, 47, 56, 58, 65]

re-evaluate their tracking tool choice during use; 4) rationales behind design choices considered in self-tracking research prototypes; and 5) empirical evaluations of the design choices of research prototypes. We further referenced the Paper Browser from Epstein et al.'s literature review of personal informatics [24] to identify papers with empirical studies examining preparation [50]. The literature we examined considered varied domains including physical activity, sleep, diet monitoring, mental health, chronic conditions, and productivity.

We used affinity diagramming to find common themes that participants expressed when evaluating the utility of self-tracking tools, which are represented as *factors* in Table 1. These factors were then grouped to form *attributes*, which represent features or factors a person might consider or want to consider when selecting a tracking tool. Our attributes draw on findings from approximately 50 relevant publications, and are roughly sorted by their prevalence and described importance.

3.2 Findings

We identified seven attributes of features and factors people assess to determine if tracking tool serves their needs: *data collected*, *feedback provided*, *goal-setting capabilities*, *privacy*, *social opportunities*, *style*, and *convenience.*

3.2.1 Data Collected. Past literature has shown that people frequently consider and reflect on what data a personal tracking tool is capable of tracking. The *data collected* attribute is formed by factors *what it collects* and *how it is collected. What it collects* describes the type of data a tracking tool supports obtaining, such as the type of food being consumed [53, 65], running activities [14], and amount of sleep one has gotten [20]. *How it is collected* is concerned with how the data are being recorded in trackers. Participants in prior studies have highlighted the importance of how the tracker acquires the data they track. For example, many systems allow people to input their data manually [17, 46], while some utilize sensors to automate the process [12, 47]. Mixed tracking methods have also been proposed [11]. Although manual data collection techniques may enable a person to obtain deeper understanding of the practices they are tracking [11], many people find the process overly burdensome [17, 25, 47]. People also consider the accuracy and completeness of the collected data [12, 17, 74], which are often influenced by how the tracking technology supports data collection [11].

3.2.2 Feedback Provided. Prior literature has often described how trackers process data to generate feedback and considered how participants perceive that feedback. The *feedback provided* attribute considers the *kind of feedback*, which describes how self-tracking tools process data when creating insights for a person. For example, tracking tools can aggregate daily or weekly totals based on the tracked data [15, 16, 26], provide summaries describing current practices [6], or suggest behaviors to undertake based on one's physical activities [32]. Once the data are processed and insights are generated, people express different preferences towards *how it's presented*. Examples include text descriptions [6, 20, 21], using calendars to allow people view their data over multiple dates [18, 27, 65, 71], and graphs or other visualizations [12, 42]. Participants in prior studies have expressed varied preferences for how technology should present feedback, often based on their level of experience with collecting data or amount of effort they wished to place on reviewing feedback [6, 12, 62].

3.2.3 Goal-setting Capabilities. Participants in prior studies have frequently considered how features supported goal setting and monitoring, especially those who sought to change their behavior. As shown in Table 1, *goal-setting capabilities* attribute are formed from factors *kind of goal*, *goal duration*, and *goal configurability*. The *kind of goal* a person would want to achieve represents the class of goals able to be set in the tracking tool, such as better sleep [21], step goals in physical activity trackers [16, 56], and body weight goals [53]. Prior literature has also discussed that people consider how frequently goals could be set or the duration of goals [34, 58] as important factors to supporting goal setting (*goal duration*). Lastly, research has suggested that some people desire technology which allows them to customize their tracking to their goals (*goal configurability*) [47, 56, 64, 65].

3.2.4 Privacy. Participants and researchers have frequently expressed *privacy* concerns around the data that tracking tools enable collecting. Prior literature has discussed that a trustful relationship between the tracking tool and the person using it contributes heavily to the overall experience [45, 58]. Thus, people are often concerned how much of the data they collect are being shared [24, 34] with other parties such as the application developer (*data disclosure*). People also consider whether they have control over the *history* of their data, such as being able to delete it after abandonment [45] or whether it persists after they examine it [26].

3.2.5 Social Opportunities. Prior literature has stressed that self-tracking practices were often social endeavors, where people would seek out support, encouragement, or advice from others. Thus, the *social opportunities* attribute is formed from factors *social features* and *what is shared*. *Social features* support social interaction, such as the ability to follow friends on social media [13, 26], creating support groups [45], and online chatting features [33]. Prior to using social features, people also consider what the features support sharing (*what is shared*), with some preferring sharing tracked activities while others wish to selectively choose accomplishments [15, 26].

3.2.6 Style. Statements from prior literature has suggested that a person's stylistic identity can also contribute to how they perceive a self-tracking tool. *Aesthetics and feel* describe stylistic properties of a self-tracking app, which includes company brands in determining whether a self-tracker is a good stylistic fit [14, 34, 47]. Examples include sporty, serious, and playful styles. These stylistic choices can follow stereotypes, such as gendered design, which can make people feel excluded by their self-tracking apps [27]. Some research has suggested that technology could aim to support customizing styles through skins or custom backgrounds [36].

3.2.7 Convenience. We grouped the factors *third-party support* and *notifications* to form the *convenience* attribute, as participants from prior literature have felt that these features generally aimed to lower the burden of use.

People frequently desire self-tracking tools that integrate with existing third-party tools, such as to support exporting tracked data [12, 45], Bluetooth support with other devices [47], and integration with social media services [26]. Prior research has found that notifications can be helpful for reminding people to track [5, 42] or suggest ways of changing behavior based on passively-collected data [32].

4 SURFACING IMPORTANT ATTRIBUTES OF SELF-TRACKING APPS IN AN APP STORE

Informed by the seven attributes uncovered from reviewing literature, we designed a low-fidelity prototype that modifies the design of typical app stores by tagging apps according to these attributes and enabling searching and filtering by the tags. We opted to make low-fidelity prototype as they can help explore the application of design ideas [8]. Furthermore, when compared to high-fidelity prototypes, lower fidelity prototypes are useful for obtaining feedback on design concepts [63].

We explored how technology could help reduce the burden of app selection, since people frequently use digital resources to identify and seek advice on tracking tools [12, 28]. We decided to investigate how attributes could be presented in app stores, as they are the default distribution points for major smartphone platforms (e.g., Android, iOS) and already include mechanisms aimed to support app selection such as descriptions, ratings, and reviews. In addition, because app stores are a common place where people learn about apps and evaluate them [41, 70], they are a natural place to include additional mechanisms which describe self-tracking apps. We opted to redesign an app store rather than explore increasing flexibility of individual apps because past work has argued that app ecosystems, rather than individual apps, are better suited to supporting people's varied self-tracking preferences and goals [55, 57]. We elaborate on opportunities for other kinds of technology to better support selection in our discussion, such as strategies for summarizing features of apps in online informational sources.

4.1 Design of a Low-Fidelity Prototype App Store

We designed a low-fidelity prototype that incorporates the attributes we surfaced from the literature review into an app store by including *tags* that display features corresponding each attribute. The prototype is organized into two pages: the *app description page* (Figure 1) and *the app store page* (Figure 2). The *app description page* is dedicated to presenting an individual self-tracking app and showing its tracking-related features via tags. The *app store page* showed all registered self-tracking apps and includes the ability to filter by tags, allowing a person to search by selecting the tags they desired using. We created the prototype with a drawing program, adding the tags of a few tracking apps as examples for what tags a fully-realized design might include. When creating the prototype, we included six existing apps that the first author had extensive experience using: three productivity apps and three fitness trackers.

4.1.1 Tags and Attributes. Because the prototype aimed to expose features related to the seven attributes of self-tracking apps we identified as potentially valuable for selection (Table 1), our prototype describes represents each feature associated with a short discernible tag. We opted for tags to describe the features composing each attribute in a simplistic manner, and to label attributes in a way which would enable a person to identify apps which had those attributes through techniques like searching or filtering. The tag-based system was presented in both the app description page and store page of the low-fidelity prototype. These tags were organized into six attributes: What it tracks (*data collected* attribute), feedback and analytics (*feedback provided* attribute), supporting goals (*goal-setting capabilities* attribute), convenience (*convenience* attribute), and privacy (*privacy* attribute). We decided not to use tags corresponding to the *Style* attribute given the attribute's subjectivity, and because we felt a person would be better-served to determine the stylistic fit themselves based on screenshots and app icons. Features such as those shown in "Examples" in Table 1 created the tags for their respective attributes.

We designed the tag-based system to be hierarchical, where high-level, more abstract tags describing a particular factor could contain lower-level, more specific tags describing how the app implemented that factor. For example,

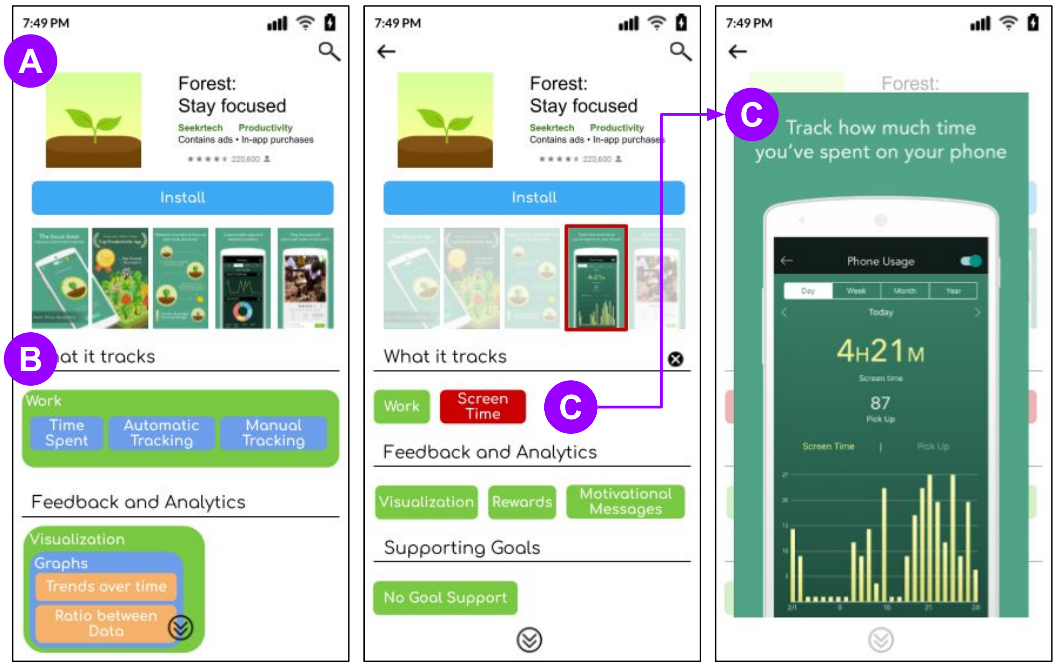

Fig. 1. The app description page of our low-fidelity prototype. (A) The page contains information typical of description pages on app stores, including an icon, screenshot, and number of reviews. (B) The page organizes tags into attributes, with sub-tags to further explain how tags are instantiated. (C) A person may select a tag to highlight screenshots showing the corresponding feature(s) in the app.

the "Work" tag from "What it tracks" could contain child tags such as "Time Spent," "Automatic Tracking," and "Manual Tracking" to indicate that the app would track a person's work by monitoring time, and provides the option to be automatic or manual. Similarly, the "Visualization" tag in "Feedback and Analytics" could contain the child tag "Graphs," which could contain child tags such as "Trends over time" or "Ratio between data" (Figure 1B).

4.1.2 App Description Page. The app description page first shows information typical of existing app stores, including the app icon, metadata (e.g., name, store category, developer, price), presence of in-app purchases and ads, average star rating, number of downloads, and screenshots of the app (Figure 1A). The prototype includes our attributes and tags in place of the developer-written description of apps, organized by the six attributes as discussed in Section 4.1.1 (Figure 1B). To help a person understand how an app instantiates the tags, the prototype allows a person to click on a tag, which then highlights the pictures that are related to that tag (Figure 1C).

4.1.3 Store Page. The store page includes elements typical of existing app stores such as a search bar to search for apps by name and app icons highlighting popular or prominent apps. The apps shown in the prototype are organized into categories based on the "what it Tracks" factor of the "Data" attribute (Figure 2D). We organized apps this way because this factor was commonly mentioned in prior literature as a reason why tracking tools did not meet a person's needs or as a barrier to understanding whether an app would support a person's goal [14, 25, 28, 36, 47]. Alongside the app icon, the prototype displays a few tags associated with the app that past literature have suggested are important attributes for selection (e.g., lower-level tags describing what an app

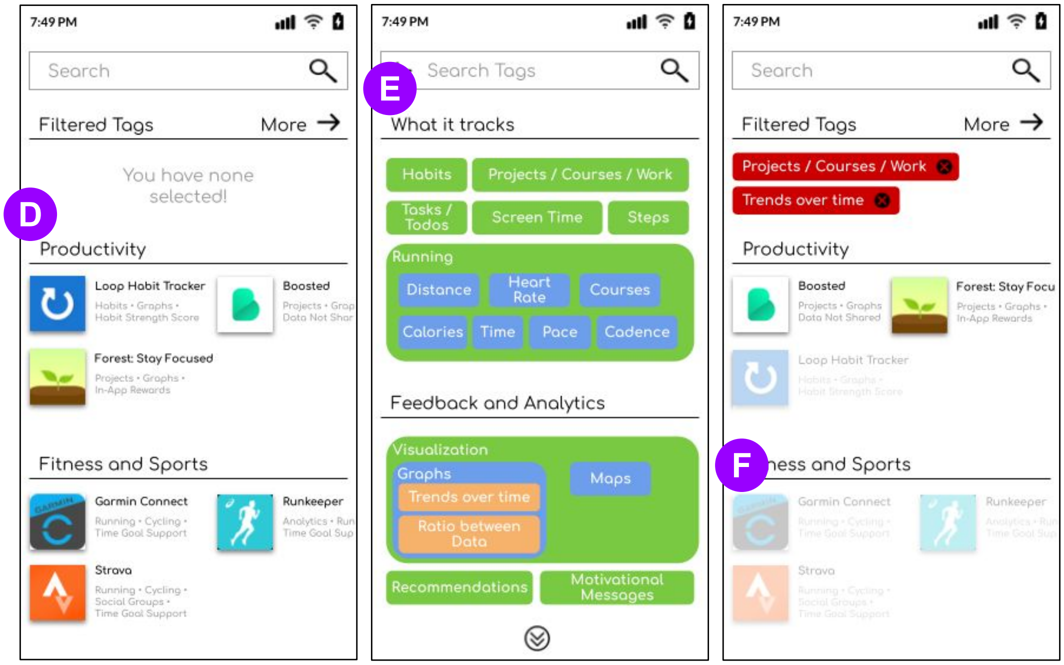

Fig. 2. The app store page of our low-fidelity prototype. (D) Apps were categorized by the data attribute, with cards highlighting the app's icon, name, and a few prominent tags. (E) A person may select tags to filter by, which upon selection, (F) de-emphasizes apps that do not fit the filter.

tracks, how feedback is presented) or tags relatively rare to apps in that category (e.g., rewards, goals support, social features). The store page allowed a person to search for apps directly using keywords, browse through categories by scrolling down, and browse through apps in each category by scrolling sideways.

As part of searching for apps, a person could use tags to filter apps to highlight those which contain all or many of the tags they find most valuable (Figure 2E). After selecting tags, the store page would then highlight the apps that includes all tags and gray out (i.e., makes transparent) apps which do not (Figure 2F).

5 INVESTIGATING PEOPLE'S PERCEPTIONS OF TECHNOLOGY SUPPORT FOR TRACKING APP SELECTION

To better understand how people approach self-tracking apps (RQ1 & RQ2) and elicit feedback on our use of attributes to understand how technology can better support selection (RQ3), we interviewed 18 participants with prior experience with self-tracking.

5.1 Method

We conducted semi-structured interviews divided into two parts. We first asked participants how they had searched for and selected the self-tracking apps they had previous experience using, and asked them to share what about the apps they considered during the process. We then guided the participant through the app description page and store page of the low-fidelity prototype to understand their perceptions on each attribute, labeling attributes as tags, and the potential utility of browsing, searching, and filtering by tags. Roughly a

Table 2. Our 18 interview participants had used and were using tracking tools in a variety of self-tracking domains. Domains that the participant is currently tracking are marked as [domain]*.

ID	Gender, Age	Profession	Tracking Domain (Prior and Present*)
P1	(F,28)	Accountant	Pet Behaviors*, Physical Activity
P2	(M,22)	Cashier	Personal Finances*, Physical Activity, Food, Weight
P3	(M,27)	Product Manager	Personal Finances*, Physical Activity, Sleep, Food, Weight
P4	(F,22)	Student	Women's Health*, Physical Activity, Sleep, Personal Finances
P5	(F,32)	Lawyer	Women's Health*, Physical Activity, Productivity, Managing Any Chronic Condition, Mental Health or Wellbeing, Sleep, Food, Weight, Personal Finances
P6	(M,47)	Programmer	Physical Activity*, Weight, Personal Finances
P7	(F,30)	Caseworker	Sleep*, Physical Activity, Food, Weight, Women's Health
P8	(F,30)	Writer	Physical Activity*, Sleep
P9	(F,24)	Teacher	Women's Health*, Physical Activity, Sleep, Food, Personal Finances
P10	(F,31)	Marketing	Productivity*, Physical Activity, Sleep, Food, Weight, Personal Finances
P11	(F,29)	Engineering Manager	Personal Finances*, Food, Weight
P12	(F,29)	Chemist	Food*, Sleep
P13	(F,27)	Data analyst	Women's Health*, Physical Activity, Productivity, Sleep, Weight, Personal Finances
P14	(F,26)	Digital Marketer, Web Designer	Mental Health*, Physical Activity, Productivity, Sleep
P15	(M,26)	PhD Student	Personal Finances*
P16	(M,38)	Scientist	Managing Any Chronic Condition*, Physical Activity, Sleep, Food, Weight
P17	(F,25)	Program Coordinator	Women's Health*, Physical Activity, Sleep, Food, Weight, Personal Finances
P18	(F,21)	Student	Physical Activity*, Sleep, Food, Weight, Women's Health

third of the interview focused on understanding participant's prior experiences selecting self-tracking apps, with the remaining time focused on describing how tags and attributes were instantiated in the prototype and eliciting feedback. The interview protocol and low-fidelity prototype were further iteratively refined by conducting four pilot studies with colleagues. Some of the changes from these studies included organizational changes (participants from the pilot studies better understood the concepts of attributes and tags when seeing app description page before the store page), adding verbal explanations of the tags and attributes in the prototype, and visual changes to the prototype like increasing text size.

We opted to structure the low-fidelity prototype portion of the interviews to be exploratory, rather than having participants use the prototype to complete a particular task, to investigate how the tags and attributes support their personal approaches to searching for and selecting self-tracking apps. We particularly wanted participants to reflect on their past experiences selecting self-tracking apps to understand if the strategies we implemented in the prototype would have helped address barriers they faced to selection. Although a task-based evaluative method could have offered deeper insight into whether the structure of our low-fidelity prototype successfully helped people identify features of tracking apps, participants would have incorporated fewer details or observations from their lived experience of trying to select self-tracking apps.

We recruited participants from online forums and social media in our local area (e.g., online community groups in Facebook, subreddits associated with nearby cities), as well as snowball sampling. We asked interested

Table 3. Features and aspects of self-tracking tools participants desired aligned with dimensions described by prior literature.

Attribute	Participant IDs	Example Quotes
Data Collected	1, 3, 4, 5, 6, 7, 8, 10, 12, 13, 14, 15, 16, 18	*P3: "The first thing I think about for the features set is does it track my location and all that, and what kind of metrics does it report."*
Feedback Provided	3, 4, 10, 11, 12, 13, 15, 16, 17, 18	*P11: "If I'm looking for something with tracking, visualization is probably the number one thing that I want because, like I said, I focus a lot on trends."*
Goal-setting Capabilities	10, 13	*P10: "a lot of the goals they have are pretty set so it wasn't as easy for me to customize what I wanted to do."*
Privacy	5, 6, 8, 11	*P5: "I wanted to know what it did with my data, especially for something like a period tracking app, that's small."*
Social Opportunities	1, 3, 6	*P1: "DogLog came up and one of the big benefits to that one was that it could be shared between multiple users which was what we were really looking for."*
Style	3, 4, 5, 10, 13, 14, 16, 17	*P10: "I would go into the screenshots and scroll through the screenshots and see how everything looks, if it's visually appealing."*
Convenience	2, 3, 5, 6, 7, 9, 15, 16	*P7: "Again, probably the reviews, looking to see what it offers like if it's compatible with my watch, and what it's tracking."*

participants to complete a screening survey, selecting participants that had prior experience searching for and using self-tracking apps. We specifically required participants to 1) be 18 or older, 2) have prior experience downloading a self-tracking app from an app store, and 3) be currently using a self-tracking app. To obtain perspectives from a diverse group, we further aimed to select participants who had tracked a variety of tracking domains and had varied occupations. Table 2 shows an overview of the participants that were selected for interviews. 13 participants identified as female and 5 as male. Their ages ranged from 21–47 ($\bar{x}$ = 28.6). Interviews were about 75 minutes long, and participants were compensated $25. We refer our participants as *P##*.

After the interviews, we used a third-party service to transcribe the interview audio. We employed thematic analysis [7] to analyze the data, first using a bottom-up approach to generate an initial codebook, then collating the codes into broader themes by grouping and refining our codes. Our final analysis included 94 codes in overarching six themes, including past experiences trying to find apps, priorities when searching for apps, and benefits and drawbacks of searching for apps by tags.

Our supplementary materials contain the full prototype, the interview protocol, and all codes and themes identified in the analysis.

5.2 Findings

We first describe similarities and differences between the aspects and features participants considered when selecting tracking apps from past literature. We then describe how participants approach the process of selecting self-tracking apps, learning that participants often experimented with self-tracking apps before deciding on what to use, but encountered barriers to determining whether the tracking apps would be useful for collection or reflection. As we describe the self-tracking app selection process, we also discuss the similarities found in existing research and highlight important differences. We finally present how participants perceived technology support for self-tracking selection by describing their perspectives on the low-fidelity app store prototype.

5.2.1 RQ1: What People Consider and Prioritize during Selection. Although participants' specific needs for apps varied somewhat by application domain, participants typically focused on features that are specific to self-tracking apps and understanding how they are implemented. All 18 participants described looking for features that pertain to at least one of the seven attributes we identified in our literature review during their previous experiences selecting self-tracking apps. Table 3 shows which participant considered which feature from the seven attributes. Participants most commonly considered the *data collected* attribute (15 participants), where they looked to see how a tracking app supported collecting data. For example, *P18* commented, *"if I were to search on Reddit, I would be like, 'Oh does anyone know of any good calorie apps that has a QR code in it?"'* Some participants also considered how accurate the collected data might be. *P3* noted that the accuracy of average movement time was a reason why he switched exercise tracking apps from Strava to Runkeeper, saying that *"[Runkeeper] had a feature that essentially made the tracking more accurate to when I was moving."* However, others described being less concerned about accuracy, such as *P11* using Renpho to track body composition metrics like BMI or bone mass even though she didn't *"100% believe that this [Renpho] very accurate."* When participants wanted to understand the style of apps (*style* attribute), they usually viewed pictures and screenshots displayed on app stores or online reviews to assess stylistic fit. *P4*, for instance, commented, *"I sometimes just download an app because I think it looks pretty, because I like pretty things so I'll just look through photos first"*.

Participants also frequently considered more general factors discussed in prior literature, such as overall usability and interaction design, visual appeal, and privacy (see Section 2.1). In our interviews, participants expressed that they considered the reputation of the app (thirteen participants), whether the app was continuously supported by the developer (six participants), the monetary cost of using the self-tracking app (nine participants), what permissions the app required (two participants), the app's size (two participants), and whether the app required people to register (one participant). When utilizing these information, participants described having an internalized threshold of which they would use when downloading self-tracking apps. *P6* explains, *"I mean, there's definitely a threshold, right? Four and a half to five is ideal. But yeah, if it's one star, that's bad, or half star or whatever. And then probably the number of downloads"*. From our results, we found that factors people considered when searching for other app categories are also applicable to self-tracking apps.

5.2.2 RQ2: How People Select Self-Tracking Apps. Participants' process of selecting self-tracking apps was similar to prior characterizations of selection processes in apps generally [1]. Participants typically searched for apps using multiple sources of information, then downloaded apps to trial before tracking and acting. When identifying potential self-tracking apps, participants described using a similar range of resources to prior work in other app selection domains [1, 41, 70].

Process of trialing self-tracking apps. Similar to selecting other types of apps, participants described trialing one or more self-tracking apps briefly, usually between a couple of days to weeks. When searching for a journaling app for tracking moods, *P14* said *"I wanted to test a couple of them because I never really used a app like this before. After about a week, I decided to go with the app that I'm currently using." P11* described identifying that You Need a Budget was not a good choice for them *"just a couple days into using it,"* and then sought out alternatives *"through a Google search". P10* similarly described identifying that they would need to try a different habit-tracking app *"maybe within two or three weeks"* of using Fabulous, seeking out recommendations from social media. One reason why the length of trialing self-tracking apps varied is the fact that self-tracking apps require some amount of accumulated data to portray insights, which are considered important features for self-tracking apps. When trialing self-tracking apps, participants also used multiple apps either simultaneously or sequentially. For example, *P14* trialed two journaling apps simultaneously before settling on Jour: *"I just was shopping around and actually downloaded two anxiety apps. So I had this app [Jour] and then I downloaded a different app called Bloom, and I was testing both of them out at the same time."* Other participants, such as *P1*, preferred evaluating tracking apps one at a time: *"I found one, thought I'd give it a try, downloaded it, didn't like it, got rid of it and searched again."*

For some participants, trialing involved tracking in ways similar to how they would track if they were planning to use the app longer-term. When *P12* began trialing MyFitnessPal, she quickly started tracking what she ate even while examining other apps since she might use the data later on: *"it was a couple of days in, I was trying to figure out if there was a way, if I was just missing the breakdown because they advertised that they had the information I needed, but it's behind a paywall essentially. So I knew pretty quickly [that I did not want to continue using it] but I just kept putting the information in anyway just in case I needed it while I found a substitute."* Others would not collect data they intended to use for self-understanding, but instead entered inconsequential or example data to understand how or whether the app supported their intended practice. For example, *P16* described trialing multiple apps by: *"just downloading [an app] up and saying, 'Okay, let me make a workout. Let me actually try out all the features to make them workouts.' If it's a food tracking app, I'd be like, 'Okay, let me see how easy it is for me to add a meal.'"' P18* similarly described a difference between trialing and using a food journaling app: *"Trying it out is just me playing around on the app and saying, 'Okay. How do I add this food item? Or how do I add this water to my daily or whatever.' And then actually actively using it, is like when I'm using my scale and trying to input the exact measurements."*

Since self-tracking apps require a reasonable amount of accumulated data to produce insights, participants were typically unable to evaluate their usefulness for reflection when trialing, and instead imagined whether they would find the reflective strategies useful. For example, when trialing budget apps, *P2* considered how some apps provided future insights from tracked data, saying *"another app that I tried, it was PocketSmith. One thing that I liked about that one is it has a lot more planning, and it can predict the future based on spending in the past."* Similarly, *P11* considered how different budgeting apps supported longer-term planning and summarization of her financials, trying a few apps until she found one which offered the kind of long-term support she desired. She said *"for me, budgeting is a lot more about just the overall trends... things like that where it's like, 'wow, we really had an outlier month in May. What happened there?'... and for me, that's just very well covered by Mint"*.

Participants were motivated to trial apps to better understand what data could be collected, how insights were delivered, and what other features might impact their ability to use it to better understand their habits. For example, *P14* described wanting a fitness app where she could specify an activity routine aligning with her fitness level, saying *"so many of the apps are trying to push their routines, and made by some guy that has ridiculous number of abs, and I'm like, 'That's not me.' ... So the app that I ended up settling on was much more flexible. It has fewer features, but at least it will let me customize it to how I want to do my routine. But it really took me downloading all those and trying them, and signing up for a bunch of free subscriptions, and then canceling them to determine which ones would not work for me." P11* quickly identified that You Need a Budget would not help support her budgeting goal, saying *"after I tried it, I realized it's more about paying off debt and more about restriction your spending and bucketing in a way that just didn't really mesh with how I was going to use a budgeting app."*

Participants often found apps supported tracking differently from how they expected based on their descriptions. *P13* commented that apps were often easier or more complex than she anticipated, saying *"sometimes an app will have like features that is not as easy to use to track specific health things. Sometimes an app will be easier to use than I had anticipated... Sleep Cycle specifically only has one feature where they measure your sleep cycle... that was something that was very easy to navigate when you use it."* A few participants felt it would be less time-intensive to try the app than to thoroughly read and comprehend the text descriptions. *P1* commented, *"instead of reading all this [descriptions] I'm going to download it and see if it's what I want." P13* felt the same way, saying, *"I know that usually when I download apps, I just skip over this section [text description] too because some apps will have very long descriptions and it doesn't really matter, that doesn't usually affect whether or not I download an app."*

Refining tracking app preferences. Participants expressed that they would learn more about their preferences for self-tracking apps as they continued to use them, both while trialing and after adoption. The preferences they learned and experiences they had would impact them the next time they decided to select a self-tracking app, either after abandoning an app they were trialing or returning to tracking again after lapsing [28]. For example, *P7* had some technical issues when trialing the sleep tracking app Pillow, but her experience with the app led her to seek out other apps that also included a *"smart alarm that the app sees what part of the sleeping cycle you're in. And if you're in light sleep, to wake you up around the time you have selected... I found it important because when I started using it with Pillow, I noticed that I would wake up less tired"*. *P5* imagined using her prior experience to specifically look for certain features in her future searches: *"going in, I didn't know what I wanted, but after having experienced one then I know what I would be looking for in a new one."* *P10* similarly described that her initial efforts to find productivity apps *"started off very general"*. She used a few different apps for weeks or months at a time, learning that the apps she tried *"didn't have a way to track the things that I wanted to track"* and *"didn't have goals either... Because it doesn't have a goal that you need to meet, I had a hard time feeling motivated enough to use it"*. She added, *"I didn't know that I had preferences and then as I continued to use more different types of apps, I figured out, I was able to narrow down what it is that I look for immediately."*

5.2.3 RQ3: Supporting Self-Tracking App Selection. Participants felt that tags surfacing capabilities of tracking apps could potentially better support them during selection. Participants thought that exposing app characteristics at a point prior to download (i.e., in the app store) could reduce the burden of trialing multiple apps and searching multiple information sources. For example, *P5* said: *"something like what is proposed here [the prototype] would be a lot more helpful for someone like me, where I can just like, 'What's there? What isn't there? How does it work?' I can just go from there, instead of this trial and error thing that I do now."* However, participants also expressed that utilizing tags and filters could inadvertently introduce barriers to selecting apps. We present participant's perspectives on the prototype around four themes: 1) filters and tags improving participant's understanding of app's capabilities; 2) challenges participants had interpreting tags; 3) supporting discovery of new tracking domains or capabilities and comparison between apps; and 4) searching or filtering by desired features.

Improving understanding of app capabilities. Participants particularly appreciated how a tag-based system could provide a brief overview of what data an app supported collecting without having to read long descriptions. For example, *P7* described, *"I don't have to go through every application [description] to read what they offer ... So I can just go to the tags and just spend less time selecting what I'm looking for."* *P8* highlighted that describing tracking features could help her identify what was most important, saying, *"then I can start prioritizing like, 'Okay, do I really need the finance app, that feature, or can I just...' It helps me prioritize what tags and features I want in the tracking app."* Similarly, *P4* imagined that privacy tags could help her to exclude apps: *"If I see a red flag regarding privacy, we'll share data no matter what, whatever that key phrase is, if I automatically see that in an app, that will discourage me from getting a specific app."* One of the reasons why participants preferred the prototype was that the attribute organization provided a unified structure. *P4* appreciated the unified structure, saying: *"each app that I click on [has] a different format of how they describe what the app does. This type of interface will force app makers to highlight what exactly is happening in the app. It's much more streamlined like a bullet point, an outline."*

Participants also valued how tags would enable them to quickly identify how feedback was supported within apps. *P5* noted, *"the feedback and analytic section tells you how it does, what it says it does, and what does it do... There's a million productivity [apps]... not all of them have the visualization. You know?"* *P12* imagined that being aware of what kind of feedback an app provided could have saved her time trialing: *"because I could say that I'm more interested in for example, I'm more interested in analytics than calorie tracking, versus my fitness Cronometer, so potentially this could have saved me the time of downloading the first two apps I tried."* Participants appreciated that a tag system could connect feature tags to screenshots to further show how apps supported tracking and feedback. *P7* explained: *"if you click on work, and screen time, it highlights the pictures above. That way, if I want to try my screen time, at least I get a picture of what it would look like, what that looks like to the app, I guess."*

Challenges interpreting tags. Although most participants felt that tags would be helpful for understanding how apps implemented self-tracking features which were important to them, participants were sometimes confused about what a tag aimed to describe. *P17* added that some tags might be hard for someone new to tracking to interpret, adding, *"I think the feedback and analytics can probably get lost to the average person if they don't know what feedback and analytics means in regards to an app."* People may have different expectations for what a tag means, or for how features should be categorized. For example, *P11* noted: *"I would say 'Tasks' and 'To-Dos' and 'Habits' are pretty close, and so distinguishing which ones I'm actually looking for may be difficult, or it may just include both of those tags just to be sure I cover the apps that I'm looking for."*

In addition, participants felt that tags would only be able to show what features the tracking app contained, and would fail to portray how the feature works. For example, *P2* commented on a distinguishing feature of tracking productive time with the *Forest* app: *"when you do work and you track your time in Forest you grow trees. And then if you don't do a good job with tracking your time, the trees die. So of course you want to have a healthy Forest. But none of that information is really conveyed in the tags. And I don't think there really would be a way to do that in the tags because that kind of thing is unique to this one app."* In the low-fidelity prototype, the concept of keeping trees alive to motivate a person to keep tracking productive time were depicted as "Work" when using a tag. The difficulties of portraying the distinguishing mechanics of certain features can also be problematic when comparing two apps with similar features. In the low-fidelity prototype, the tags for *Boosted* and *Stay Focused* apps are similar due to the fact that both are for tracking productive time. *P8* expressed concern on how she would compare the two apps, and said *"a lot of them, like say Boosted and Stay Focused, I see that it's pretty similar ... how do I compare these two apps if the descriptions are the same."* She later commented that she would *"have to judge it by how pretty the picture looks."*

Supporting discovery and comparison. Some participants expressed that tags could help them learn about capabilities of apps that they would not have known about or expected. For example, *P12* described that the tag for *social features* introduced her to the idea that these could be included in an app: *"I like the social features option because I wouldn't necessarily expect, for example, the one that's here, a focus tracker to have a social component to it, so that's nice to see too... of course I'll be looking forward to see if there's any additional features to this app that I am interested in."* *P14* similarly expressed learning that tracking apps could provide rewards or badges: *"I'm talking about something that I didn't know I wanted this. I didn't know I wanted to have rewards within an app, but it just opens me up to different features."* *P2* remarked that tags might encourage him to reflect more on apps' privacy: *"privacy is also honestly usually not something I think so much about when I'm choosing an app, but having that there I do like it, even though it may not be my primary consideration when choosing an app, but I think it's good to know."*

A few participants suggested that an app store organized by what data tracking apps enable collecting could introduce them to other domains they had not considered tracking. For example, *P8* described, *"if I initially went in looking for a fitness app, and then I see screen time. It was like, Oh, that looks okay. Let me put that in too". P13* similarly commented, *"I think it would encourage me to explore other types of apps and to just search for what I'm looking for because I don't necessarily look at apps based on for instance, their ability to keep track of time".*

Given that participants frequently identified multiple apps which might suit their self-tracking needs, they valued how tags might more easily support comparison between apps than the descriptions in traditional app stores. Participants imagined using tags while quickly scanning apps to compare what each supported. *P8* said, *"I like the descriptions, like these texts will say running, cycling. I like those texts. It gives a very quick overview of what it tracks compared to other apps. I like the single text here because I'm comparing different apps against each other." P16* similarly imagined comparing what data different apps supported collecting: *"it would be easier to flip through each app and see... compare the tags or compare these features, and, 'Oh, this one has a really convenient feature, that I haven't thought about. This one tracks... Like One Drop tracks a lot more data than this other one,' and* *that's really useful." P5* imagined the benefit of having *"a side by side comparison of apps that do similar things. For example, I could look at Boosted versus Forest and then I only have to compare the two apps instead of looking through a bunch of them and being like, 'This is trash. This isn't what I'm looking for.' If I wanted to compare it to similar apps, I would click on a particular tag and then look at the apps that have that particular tag."*

However, participants felt that tags could be overwhelming if they were not of interest to them or did not help with comparing apps. For example, *P6 "don't find it [social features] all that interesting"* when selecting running apps. While participants generally valued understanding what an app tracked and how it presented feedback, what other attributes participants considered important varied, feeling that including more tags compounded to the complexity of the information being presented. *P4* commented that *"when an app store gives me too much information, I feel like it overwhelms me and I just scroll past." P16* shared the same sentiment, noting: *"I could just imagine getting overwhelmed with too many bits of information. Which again, I understand in some ways, it's useful. In some ways, it's, 'oh, this is, this is so much, so many things.'"*

Searching and filtering by features. Participants appreciated how having apps labeled by features could enable filtering or searching by those features. Fourteen participants expressed that filtering by tags could help them refine their searches when trying to select an app. *P5* described: *"this [prototype] helps me be like, 'I am looking for these three things. Does this app do what I want it to do, yes or no?'" P7* shared the same sentiment, adding: *"instead of going through every single one and seeing what they are like, you can filter from what exactly you need, and then just focus on that."* A few participants noted that the ability to search by tags could alleviate the need to trial as many apps or trial apps at all. *P17* expressed that filtering by features can help her save time searching for self-tracking apps, where she commented: *"I think that if I was looking for a particular thing, like something on Lose It! that I like, or it's a caloric tracker or something like that, then I could do that and I could wipe out other apps. ... [tag filtering] really eliminates the time of installing something, finding out that I don't actually like it or it doesn't do what I want it to do, and then uninstalling it and going on the search again." P18* described how filtering could save her from having to ask for recommendations from online forums: *"if I were to search on Reddit, I would be like, 'Oh does anyone know of any good calorie apps that has a QR code in it?' That thing. I would probably have criterias that I would be looking for. That's why I would like the filter tag options, because then it would simplify. I wouldn't have to go on Reddit basically."*

P1 highlighted that their tracking domain often led to apps which were not relevant to tracking: *"when you put in dog tracking for example, there's a lot of stuff that's not relevant to what I'm looking for, it's all about dog training. But then it's just a list of different ones, if I could just break it up into categories or if I could narrow it down once I typed in a search phrase that would be more helpful."* She elaborated that even once she identified tracking apps, she still struggled to identify ones which included features she desired, and therefore appreciated how tags could be used to create filters. She said, *"when I searched for 'dog tracking,' there's a lot of things you could track around a dog and not all of them are what I'm looking for... if I could narrow down based on a few tags and one that had a social feature that you could have a shared profile with someone that would have made things easier because I had to resort to googling it, I think I looked at Reddit, I looked at some blogs to find what I needed."*

6 DISCUSSION

Our findings highlight that when deciding whether to use a self-tracking app, people want to understand what an app tracks, how it supports tracking, and what feedback it provides. We further found that technology exposing the attributes that people aim to understand and allowing them to search by such attributes has the potential to be helpful during selection. In this section, we discuss how our findings extend prior research on tool selection, consider how technology could better support trialing and future recommendations, and describe the limitations of work and opportunity for future research.

6.1 Understanding Tracking Tool Selection

Past models of personal informatics have described the act of selecting self-tracking tools as separate from the tracking & acting process of collecting, integrating, and reflecting on data [28, 50]. Some participants viewed selection as a distinct precursor to tracking and acting. When these participants trialed apps, they often stopped short of reflecting on the data they collected or even intentionally refrained from collecting the kind of data they planned to track. Instead, some participants logged data which resembled what they might attempt to log in the future, but was still nonetheless an example. Past research has highlighted that people often lapse in tracking or abandon the practice upon realizing that a tool is not supporting their needs [14, 28, 47]. Our findings extend this to suggest that some people intentionally select apps, sometimes multiple, that they intend not to adopt or even use for outright collection, instead aiming to understand the features they enable and using this process to better understand their own goals for self-tracking.

Furthermore, participants reported more complex selection processes that were not consistently constrained to specific stages of self-tracking. Some participants indicated that they continued to judge whether their selected app was effectively addressing their needs throughout their process of using it. For example, participants described seeking out reviews of other apps while using one to decide whether it was a better choice. While previous work has suggested that people may begin to feel locked into a tracking tool after beginning to track in order to preserve data [28, 50], our findings suggest that many people are able overcome these barriers to select more ideal tools. In these cases, people may transition to new tools without lapsing, and may continue with the process of collecting, integrating, and reflecting on data while simultaneously selecting a new tool. These findings are consistent with Rooksby et al.'s characterization of lived informatics [62], which discusses how people often interweave multiple self-tracking tools during their tracking journey. We similarly find that people interweave the use of self-tracking apps during selection specifically. Multiple participants reported trialing multiple apps when deciding which would be the best fit for them. Many lacked definitive criteria going into that selection process, developing their preferences as they learned about the tools they trialed. We also find that some people intentionally select apps which they know are less ideal for their self-tracking collection or reflection needs, but are beneficial in other ways (e.g., have better overall user experience or visual design, are more useful for social support). Aligning with lived informatics, our findings highlight that people do not always act as rational scientists when selecting self-tracking apps. A seemingly-logical direction for future work may have been for technology to recommend self-tracking apps to people based on their needs and goals. However, our findings suggest that recommendations will often fail to meet the everyday needs of self-tracking, and people may still prefer to try apps for themselves. We see more opportunity for technology to help showcase the features and properties of self-tracking apps to promote better discovery of tools to help people assess whether they fill align with their everyday experiences.

6.2 Recommendations for Self-Tracking App Design

Since people interweave several tracking apps during trialing and the distinction between trialing and actual tracking is often blurred, we echo suggestions from past literature and recommend that self-tracking tools support seamless movement of data between different apps [24, 28, 29, 61] by creating public APIs and adhering to open standards [29]. Future tools could leverage apps' APIs to allow people to compare data measurements collected by different apps, such as nutrient measurements for lookups of similar foods or step counts from a day's walking. Some people already make these comparisons manually [12, 31], so supporting comparison in a shared backend could remove tension between trialing and tracking. Further extending these ideas, research could examine how to separate selection and trialing from later stages of self-tracking by integrating reflection on data through visualization or other means inside of a shared backend. Doing so would allow people to more easily trial apps sequentially, switch apps, or select apps, but would require deeper consideration of how to support people's varied and complex needs for getting feedback from the aggregated data.

We also recommend developers and designers of self-tracking apps implement more accessible app demos or trial subscriptions that allows an individual to experience all of its features. Participants expressed frustration when a feature they sought to try out was inaccessible in a trial, with most seeking out an alternative instead of purchasing a subscription. While including all of the features an app provides may not be desirable or possible for app developers, self-trackers would benefit from further explanation of how excluded features operate to aid their selection decisions. Future work could examine how self-tracking technology could preview the full cycle of tracking through interactive demos, including how data are collected and how feedback is provided. For example, a demo of a food journaling app could allow a person to look up a food of their choice in their food database, and then later project that food alongside some default data into the visualization the app provides for reflection on progress towards a weekly calorie budget. This kind of previewing can particularly help people trial feedback features of apps which require more than one data entry or require data from multiple days, as these can be particularly burdensome for people to learn about through trialing.

It is important to acknowledge that easily supporting switching apps and providing extensive support for trialing be antithetical to the business models of many self-tracking apps. For many self-tracking apps, profits often rely on keeping people in the app where their data has been collected and leveraging brand recognition for establishing initial interest. Although it is important to continue to advocate for this flexibility, we see designing technology to better support initial selection of self-tracking tools, such as by improving app distribution platforms, as perhaps a more practical opportunity.

6.3 Recommendations for App Distribution Platform Design

Participant perspectives on surfacing tags and attributes suggest that technology can better support selection by prominently surfacing important features and properties of self-tracking apps which enables people to filter and search by said features and properties. Our findings suggest that, although there are many resources where this information can be shared, app distribution platforms like stores are a potentially useful place since stores already aggregate other important properties of apps people consider during selection (e.g., reviews, cost) and are the point of app download and/or purchase.

Our findings indicate that consistently tagging and categorizing all self-tracking apps in distribution platforms with aspects people consider when selecting self-tracking tools could enable people to more easily identify potentially useful apps and support comparison between multiple apps. One example of a similar change is Apple's recent adoption of a privacy subsection on their app store, highlighting how apps use different kinds of data they collect [2]. However, questions around how tags are being assigned would be an important thought, and trust between app developers and people selecting apps would be a challenge. In an open marketplace, developers may be incentivized to deceive users of self-tracking apps by indicating that their app supports features it does not, or minimizing potentially concerning features such as data disclosures. In developing our prototype, we identified the tags of the apps our prototype described based on our prior use of the apps. Similar models could be adopted to allow experts to review or endorse apps. For example, a dedicated third-party of domain experts can manually create tags or the developer themselves can include tags when registering their app in the app store. Past research has also suggested that tags can be extracted by analyzing frequently-used words in app store reviews, which often describe the most useful or unique features of an app [51, 75].

To better support self-tracking app selection, our work suggests that app distribution platforms could implement mechanisms for searching and browsing self-tracking apps by features. One strategy for implementing this tagging system is to introduce ratings for categories of tags or attributes, similar to how online resources rate broader categories of apps like their overall user experience [57]. For example, a person interested in self-tracking apps that provide physician-approved sleep recommendations could first filter apps by searching for keyword "physician-approved" and trial apps that have 4/5 ratings for such features. Although having a way to look at potentially interesting tags can lead to discovery, our findings suggest that inundating people with all features tagged can lead them to feel overwhelmed by the prospect of selection. Recommending tags based on the criteria people typically find most important (e.g., data collected, feedback provided) or personalizing based on people's prior search history could help reduce these burdens.

Although our work primarily examined strategies for improving self-tracking selection in app stores, we expect that important aspects of apps could be surfaced in other resources for selection such as online guides. For example, guides could use similar tag systems to emphasize the similarities and differences among recommended apps. Online resources could further curate collections of self-tracking apps which share properties that a person might prioritize during selection, such as apps which allow for data exporting, do not share data with developers, or support automatic data collection.

6.4 Limitations and Future Work

The seven attributes we identified when reviewing prior literature are likely not exhaustive of the aspects people prioritize when selecting self-tracking apps. For example, many participants in our study commented that they would consider technical aspects of the apps such as app size and permissions during selection. Although participants did not describe privacy concerns with how developers might disclose data to third party companies such as advertisers, the growing attention to data disclosure in personal informatics at large suggests that it may increasingly become a priority [24]. The attributes do not cover other factors which prior work have suggested influence app selection in general, such as the overall reputation of apps via star ratings or reviews [1, 22], broader usability categories like adherence to user experience guidelines [57], or pricing models (e.g., free, one-time purchase, subscription). Our attributes also focus on how people assess apps' functionality, but brand recognition and marketing also play critical roles in influencing selection. Although our work points to opportunity for technology to better surface the capabilities of self-tracking apps, we acknowledge the importance of these other influences on how people approach selection.

People commonly use tools for tracking besides apps, including physical devices, desktop applications, and paper journals [12, 24]. The selection processes practiced by our participants and the approach to integrating attributes into technology facilitating selection may be limited to supporting selection of phone apps for tracking, and may be different from people selecting physical devices or desktop applications. For example, the burdens to trying out other kinds of tracking may be higher due to added costs (e.g., purchasing one or more physical devices) or the demands of tracking (e.g., creating and adhering to tracking on paper). Future work is needed to understand how people approach selection of dedicated devices for tracking, and how technology can better incorporate the attributes of self-tracking tools people prioritize into other resources people leverage when seeking out dedicated devices (e.g., online review sites, online retailers, physical retailers).

Because we aimed to learn from the selection practices of people who had previously used self-tracking tools, our findings do not incorporate the perspectives who might be considering self-tracking, but have not approached selecting a self-tracking app. Although a few participants reflected on their experiences selecting their first tracking apps, we expect that newcomers to self-tracking might have different priorities or selection strategies. For example, first-time self-trackers may have less of an understanding of what insights self-tracking tools are able to provide, or the impacts of automated versus manual tracking on their everyday experiences [61]. Examining the self-tracking needs and selection strategies people that are not experienced with collecting situated data could provide valuable insight into how to support people in selecting their first self-tracking tool.

To facilitate accurate tagging, we populated the low-fidelity prototype with apps we had extensive experience using. When reviewing the low-fidelity prototype, participants were therefore reviewing apps that may have contained tags they were not interested in and not have been in a domain they were interested in tracking. Instead, participant feedback focused on the overall approach of browsing, searching, and filtering apps through tags grouped by attributes. While this feedback was valuable for seeking participant perspectives on the overall concept of selecting apps by attributes, participants may prioritize different tags or feel differently about the concept if they were to search for domains they were interested in tracking themselves. Additionally, a few participants had firsthand experience with some of the tools included in our prototype, which may have allowed them to infer the meaning behind the tags. Future research seeking to deeply understand how people prioritize different tags or categories of tags could focus on apps which participants do not have prior experience with, or incorporate hypothetical apps which fill in different niches not supported by commercial apps.

Although we selected participants who had experience selecting self-tracking apps from app stores, none of our participants were actively interested in selecting an app to fulfill a particular self-tracking need. Participants therefore imagined how they might use tags to inform their selection based on their past experiences searching for apps. How people utilize the tags and attributes in practice might therefore differ from how participants envisioned the process when interacting with the prototype. For example, people often spend under two minutes searching for apps in app stores and rarely scroll past the top five apps [22], whereas our participants spent longer reflecting on the attributes in the prototype.

Our participants skewed WEIRD [37] (Western, Educated, Industrialized, Rich, and Democratic) and younger, suggesting that their experiences selecting apps and preferences for technology support may differ from other populations. This skew is likely due to our recruitment method, where we mainly utilized online recruiting due to the COVID-19 pandemic. For example, our participants were typically adept at using technology to search for apps given their occupations and frequent use of online communities and social media, similar to other studies of early adopters of self-tracking apps [12]. People who rely less on technology may value different kinds of support for selection, digital or not. Further work is necessary to understand how people from different demographics, socioeconomic levels, and other backgrounds approach selection and desire technology support, and care should be taken when applying our findings to other groups.

Although our study only explored how app stores can better support self-tracking app selection, there are opportunities for researchers and designers to better support self-tracking selection without requiring changes to the design of app store platforms. Participants frequently made use of additional resources when making choices about what apps to use (Section 5.2.2), suggesting that standalone websites which surface and enable searching for desirable features could help people in their selection process. Sites like PsyberGuide for mental health have aimed to evaluate the overall quality and user experience of self-tracking apps [57]. Future research could explore how to design such pages to integrate how apps implement those features. Participants also suggested that current descriptions of self-tracking apps and screenshots included in app stores are often insufficient for them to understand how important features, such as how data are collected or what representations of data are provided. Future work could create standardized guidelines for what developers of self-tracking apps should aim to highlight beyond general recommendations provided by app platform developers. Participants found description consistency helpful when comparing self-tracking apps, suggesting potential value in formulating guidelines. Standalone resources could be a better alternative, since implementing the prototype presented in this work requires extensive modification of existing app stores. However, this method adds an additional medium for people to look at before deciding on a tool, and requires a person to seek out the information on their own.

7 CONCLUSION

We found that people frequently sought to understand and considered how apps support tracking when selecting a self-tracking app, alongside more general factors frequently surfaced in app stores and online reviews. To support people's practice of trialing self-tracking apps, we echo suggestions from prior work to support data portability and suggest designing demos or trial subscriptions to describe how tracking processes are implemented. While app stores and online guides can be useful aids for understanding the functionality of self-tracking apps, we demonstrate opportunity for these resources to better support comparing self-tracking apps and searching for apps by features and properties of them. Addressing early barriers to selecting self-tracking technology can prevent cascading consequences to help people successfully reap the benefits of using personal informatics tools.

ACKNOWLEDGMENTS

We thank Lucas M. Silva, Xi Lu, Seo Yeon Hwang, and Minu Park for their helpful comments in our pilot studies, and our anonymous reviewers for their insightful feedback. This work was funded in part by the National Science Foundation under award IIS-1850389.

REFERENCES

[1] Nora Alkaldi and Karen Renaud. 2016. Why do people adopt, or reject, smartphone password managers?. In *EuroUSEC '16*. Internet Society. https://doi.org/10.14722/eurousec.2016.23011 1st European Workshop on Usable Security, EuroUSEC 2016 ; Conference date: 18-07-2016 Through 18-07-2016.

[2] Inc Apple. 2021. About privacy information on the App Store and the choices you have to control your data. Online. https://support.apple.com/en-us/HT211970

[3] 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. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376809

[4] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L. Cox. 2018. *Flexible and Mindful Self-Tracking: Design Implications from Paper Bullet Journals*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173602

[5] Frank Bentley and Konrad Tollmar. 2013. *The Power of Mobile Notifications to Increase Wellbeing Logging Behavior*. Association for Computing Machinery, New York, NY, USA, 1095–1098. https://doi.org/10.1145/2470654.2466140

[6] Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *ACM Trans. Comput.-Hum. Interact.* 20, 5, Article 30 (Nov. 2013), 27 pages. https://doi.org/10.1145/2503823

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

[8] Bill Buxton. 2010. *Sketching user experiences: getting the design right and the right design.* Morgan kaufmann.

[9] Ting-Ray Chang, Eija Kaasinen, and Kirsikka Kaipainen. 2012. What influences users' decisions to take apps into use?: a framework for evaluating persuasive and engaging design in mobile Apps for well-being. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia - MUM '12*. ACM Press, Ulm, Germany, 1. https://doi.org/10.1145/2406367.2406370

[10] Jaehee Cho, Margaret M. Quinlan, Dongjin Park, and Ghee-Young Noh. 2014. Determinants of Adoption of Smartphone Health Apps among College Students. *American Journal of Health Behavior* 38, 6 (Nov. 2014), 860–870. https://doi.org/10.5993/AJHB.38.6.8

[11] E. K. Choe, S. Abdullah, M. Rabbi, E. Thomaz, D. A. Epstein, F. Cordeiro, M. Kay, G. D. Abowd, T. Choudhury, J. Fogarty, B. Lee, M. Matthews, and J. A. Kientz. 2017. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing* 16, 1 (2017), 74–84. https://doi.org/10.1109/MPRV.2017.18

[12] 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* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1143–1152. https://doi.org/10.1145/2556288.2557372

[13] Chia-Fang Chung, Nanna Gorm, Irina A. Shklovski, and Sean Munson. 2017. *Finding the Right Fit: Understanding Health Tracking in Workplace Wellness Programs*. Association for Computing Machinery, New York, NY, USA, 4875–4886. https://doi.org/10.1145/3025453.3025510

[14] James Clawson, Jessica A. Pater, Andrew D. Miller, Elizabeth D. Mynatt, and Lena Mamykina. 2015. No Longer Wearing: Investigating the Abandonment of Personal Health-Tracking Technologies on Craigslist. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 647–658. https://doi.org/10.1145/2750858.2807554

[15] Sunny Consolvo, Katherine Everitt, Ian Smith, and James A. Landay. 2006. Design Requirements for Technologies That Encourage Physical Activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada) *(CHI '06)*. Association for Computing Machinery, New York, NY, USA, 457–466. https://doi.org/10.1145/1124772.1124840

[16] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. 2008. Activity Sensing in the Wild: A Field Trial of Ubifit Garden.

In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1797–1806. https://doi.org/10.1145/1357054.1357335

[17] Felicia Cordeiro, Daniel A. Epstein, Edison Thomaz, Elizabeth Bales, Arvind K. Jagannathan, Gregory D. Abowd, and James Fogarty. 2015. Barriers and Negative Nudges: Exploring Challenges in Food Journaling. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. ACM, New York, NY, USA, 1159–1162. https://doi.org/10.1145/2702123.2702155

[18] Mayara Costa Figueiredo, Clara Caldeira, Elizabeth Victoria Eikey, Melissa Mazmanian, and Yunan Chen. 2018. Engaging with health data: The interplay between self-tracking activities and emotions in fertility struggles. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–20.

[19] Nediyana Daskalova, Karthik Desingh, Alexandra Papoutsaki, Diane Schulze, Han Sha, and Jeff Huang. 2017. Lessons Learned from Two Cohorts of Personal Informatics Self-Experiments. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 46 (Sept. 2017), 22 pages. https://doi.org/10.1145/3130911

[20] Nediyana Daskalova, Danaë Metaxa-Kakavouli, Adrienne Tran, Nicole Nugent, Julie Boergers, John McGeary, and Jeff Huang. 2016. SleepCoacher: A Personalized Automated Self-Experimentation System for Sleep Recommendations. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 347–358. https://doi.org/10.1145/2984511.2984534

[21] Nediyana Daskalova, Jina Yoon, Yibing Wang, Cintia Araujo, Guillermo Beltran, Nicole Nugent, John McGeary, Joseph Jay Williams, and Jeff Huang. 2020. SleepBandits: Guided Flexible Self-Experiments for Sleep. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376584

[22] Leyla Dogruel, Sven Joeckel, and Nicholas D. Bowman. 2015. Choosing the right app: An exploratory perspective on heuristic decision processes for smartphone app selection. *Mobile Media & Communication* 3, 1 (Jan. 2015), 125–144. https://doi.org/10.1177/2050157914557509

[23] 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*. Association for Computing Machinery, New York, NY, USA, 642–654. https://doi.org/10.1145/3025453.3025591

[24] 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

[25] 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*. Association for Computing Machinery, New York, NY, USA, 1109–1113. https://doi.org/10.1145/2858036.2858045

[26] Daniel A. Epstein, Bradley H. Jacobson, Elizabeth Bales, David W. McDonald, and Sean A. Munson. 2015. From "Nobody Cares" to "Way to Go!": A Design Framework for Social Sharing in Personal Informatics. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1622–1636. https://doi.org/10.1145/2675133.2675135

[27] 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)*. Association for Computing Machinery, New York, NY, USA, 6876–6888. https://doi.org/10.1145/3025453.3025635

[28] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 731–742. https://doi.org/10.1145/2750858.2804250

[29] Deborah Estrin and Ida Sim. 2010. Open mHealth architecture: an engine for health care innovation. *Science* 330, 6005 (2010), 759–760.

[30] Denzil Ferreira, Vassilis Kostakos, Alastair R. Beresford, Janne Lindqvist, and Anind K. Dey. 2015. Securacy: an empirical investigation of Android applications' network usage, privacy and security. In *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. ACM, New York New York, 1–11. https://doi.org/10.1145/2766498.2766506

[31] Thomas Fritz, Elaine M. Huang, Gail C. Murphy, and Thomas Zimmermann. 2014. Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 487–496. https://doi.org/10.1145/2556288.2557383

[32] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2015. How Do We Engage with Activity Trackers? A Longitudinal Study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 1305–1316. https://doi.org/10.1145/2750858.2804290

[33] Xinning Gui, Yu Chen, Clara Caldeira, Dan Xiao, and Yunan Chen. 2017. *When Fitness Meets Social Networks: Investigating Fitness Tracking and Social Practices on WeRun*. Association for Computing Machinery, New York, NY, USA, 1647–1659. https://doi.org/10.1145/3025453.3025654

[34] 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* (Brisbane, QLD, Australia) *(DIS '16)*. Association for Computing Machinery, New York, NY, USA, 286–300. https://doi.org/10.1145/2901790.2901803

[35] Mark A. Harris, Robert Brookshire, and Amita Goyal Chin. 2016. Identifying factors influencing consumers' intent to install mobile applications. *International Journal of Information Management* 36, 3 (June 2016), 441–450. https://doi.org/10.1016/j.ijinfomgt.2016.02.004

[36] Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and Jon Bird. 2015. Activity Tracking: Barriers, Workarounds and Customisation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 617–621. https://doi.org/10.1145/2750858.2805832

[37] Joseph Henrich, Steven J. Heine, and Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33, 2-3 (June 2010), 61–83. https://doi.org/10.1017/s0140525x0999152x

[38] Guei-Hua Huang and Nikolaos Korfiatis. 2015. Trying Before Buying: The Moderating Role of Online Reviews in Trial Attitude Formation Toward Mobile Applications. *International Journal of Electronic Commerce* 19, 4 (Oct. 2015), 77–111. https://doi.org/10.1080/10864415.2015.1029359

[39] Hsiao-Ying Huang and Masooda Bashir. 2017. Users' Adoption of Mental Health Apps: Examining the Impact of Information Cues. *JMIR mHealth and uHealth* 5, 6 (June 2017), e83. https://doi.org/10.2196/mhealth.6827

[40] IQVIA. 2017. The Growing Value of Digital Health. Online. https://www.iqvia.com/insights/the-iqvia-institute/reports/the-growing-value-of-digital-health

[41] Shaheen Kanthawala, Eunsin Joo, Anastasia Kononova, Wei Peng, and Shelia Cotten. 2019. Folk theorizing the quality and credibility of health apps. *Mobile Media & Communication* 7, 2 (May 2019), 175–194. https://doi.org/10.1177/2050157918796859

[42] 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

[43] Patrick Gage Kelley, Lorrie Faith Cranor, and Norman Sadeh. 2013. *Privacy as Part of the App Decision-Making Process*. Association for Computing Machinery, New York, NY, USA, 3393–3402. https://doi.org/10.1145/2470654.2466466

[44] Deanna S. Kempf and Robert E. Smith. 1998. Consumer Processing of Product Trial and the Influence of Prior Advertising: A Structural Modeling Approach. *Journal of Marketing Research* 35, 3 (Aug. 1998), 325. https://doi.org/10.2307/3152031

[45] Da-jung Kim, Yeoreum Lee, Saeyoung Rho, and Youn-kyung Lim. 2016. *Design Opportunities in Three Stages of Relationship Development between Users and Self-Tracking Devices*. Association for Computing Machinery, New York, NY, USA, 699–703. https://doi.org/10.1145/2858036.2858148

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

[47] Amanda Lazar, Christian Koehler, Theresa Jean 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* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 635–646. https://doi.org/10.1145/2750858.2804288

[48] 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* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 743–754. https://doi.org/10.1145/2750858.2807552

[49] Huoran Li, Xuan Lu, Xuanzhe Liu, Tao Xie, Kaigui Bian, Felix Xiaozhu Lin, Qiaozhu Mei, and Feng Feng. 2015. Characterizing Smartphone Usage Patterns from Millions of Android Users. In *Proceedings of the 2015 Internet Measurement Conference*. ACM, Tokyo Japan, 459–472. https://doi.org/10.1145/2815675.2815686

[50] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. *A Stage-Based Model of Personal Informatics Systems*. Association for Computing Machinery, New York, NY, USA, 557–566. https://doi.org/10.1145/1753326.1753409

[51] Jovian Lin, Kazunari Sugiyama, Min-Yen Kan, and Tat-Seng Chua. 2014. New and Improved: Modeling Versions to Improve App Recommendation. In *Proceedings of the 37th International ACM SIGIR Conference on Research & Development in Information Retrieval* (Gold Coast, Queensland, Australia) *(SIGIR '14)*. Association for Computing Machinery, New York, NY, USA, 647–656. https://doi.org/10.1145/2600428.2609560

[52] James J Lin, Lena Mamykina, Silvia Lindtner, Gregory Delajoux, and Henry B Strub. 2006. Fish'n'Steps: Encouraging physical activity with an interactive computer game. In *International conference on ubiquitous computing*. Springer, 261–278.

[53] 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* (Glasgow, Scotland Uk)

(CHI '19). Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300822
[54] Aashish Mehra, Justin Paul, and Rahul Pratap Singh Kaurav. 2020. Determinants of mobile apps adoption among young adults: theoretical extension and analysis. *Journal of Marketing Communications* (Feb. 2020), 1–29. https://doi.org/10.1080/13527266.2020.1725780
[55] David C Mohr, Kathryn Noth Tomasino, Emily G Lattie, Hannah L Palac, Mary J Kwasny, Kenneth Weingardt, Chris J Karr, Susan M Kaiser, Rebecca C Rossom, Leland R Bardsley, et al. 2017. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. *Journal of medical Internet research* 19, 1 (2017), e10.
[56] Sean Munson and Sunny Consolvo. 2012. Exploring Goal-setting, Rewards, Self-monitoring, and Sharing to Motivate Physical Activity. IEEE. https://doi.org/10.4108/icst.pervasivehealth.2012.248691
[57] Martha Neary and Stephen M Schueller. 2018. State of the field of mental health apps. *Cognitive and Behavioral Practice* 25, 4 (2018), 531–537.
[58] Jasmin Niess and Paweł W Woźniak. 2018. Supporting Meaningful Personal Fitness: The Tracker Goal Evolution Model. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173745
[59] Iryna Pentina, Lixuan Zhang, Hatem Bata, and Ying Chen. 2016. Exploring privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison. *Computers in Human Behavior* 65 (Dec. 2016), 409–419. https://doi.org/10.1016/j.chb.2016.09.005
[60] Thanasis Petsas, Antonis Papadogiannakis, Michalis Polychronakis, Evangelos P. Markatos, and Thomas Karagiannis. 2013. Rise of the planet of the apps: a systematic study of the mobile app ecosystem. In *Proceedings of the 2013 conference on Internet measurement conference*. ACM, Barcelona Spain, 277–290. https://doi.org/10.1145/2504730.2504749
[61] Amon Rapp and Federica Cena. 2016. Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human-Computer Studies* 94 (2016), 1–17.
[62] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal Tracking as Lived Informatics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1163–1172. https://doi.org/10.1145/2556288.2557039
[63] Jim Rudd, Ken Stern, and Scott Isensee. 1996. Low vs. High-Fidelity Prototyping Debate. *Interactions* 3, 1 (Jan. 1996), 76–85. https://doi.org/10.1145/223500.223514
[64] Jessica Schroeder, Chia-Fang Chung, Daniel A. Epstein, Ravi Karkar, Adele Parsons, Natalia Murinova, James Fogarty, and Sean A. Munson. 2018. Examining Self-Tracking by People with Migraine: Goals, Needs, and Opportunities in a Chronic Health Condition. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 135–148. https://doi.org/10.1145/3196709.3196738
[65] Jessica Schroeder, Ravi Karkar, Natalia Murinova, James Fogarty, and Sean A. Munson. 2019. Examining Opportunities for Goal-Directed Self-Tracking to Support Chronic Condition Management. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 4, Article 151 (Dec. 2019), 26 pages. https://doi.org/10.1145/3369809
[66] Stephen M Schueller, Martha Neary, Kristen O'Loughlin, and Elizabeth C Adkins. 2018. Discovery of and interest in health apps among those with mental health needs: survey and focus group study. *Journal of medical Internet research* 20, 6 (2018), e10141. Publisher: JMIR Publications Inc., Toronto, Canada.
[67] Patrick C Shih, Kyungsik Han, Erika Shehan Poole, Mary Beth Rosson, and John M Carroll. 2015. Use and Adoption Challenges of Wearable Activity Trackers. *IConference 2015 proceedings* (2015). http://hdl.handle.net/2142/73649
[68] Statista. 2017. Percentage of mobile medical application categories used by U.S. adults at least once as of 2017. Online. https://www.statista.com/statistics/378850/top-mobile-health-application-categories-used-by-us-consumers/
[69] Statista. 2020. Number of health and fitness app users in the United States from 2018 to 2022 (in millions). Online. https://www.statista.com/statistics/1154994/number-us-fitness-health-app-users/
[70] Besiki Stvilia and Wonchan Choi. 2015. Mobile wellness application-seeking behavior by college students—An exploratory study. *Library & Information Science Research* 37, 3 (July 2015), 201–208. https://doi.org/10.1016/j.lisr.2015.04.007
[71] Lie Ming Tang and Judy Kay. 2017. Harnessing long term physical activity data—How long-term trackers use data and how an adherence-based interface supports new insights. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 1–28.
[72] Qiang Xu, Jeffrey Erman, Alexandre Gerber, Zhuoqing Mao, Jeffrey Pang, and Shobha Venkataraman. 2011. Identifying diverse usage behaviors of smartphone apps. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference - IMC '11*. ACM Press, Berlin, Germany, 329. https://doi.org/10.1145/2068816.2068847
[73] Runhua Xu, Remo Manuel Frey, Elgar Fleisch, and Alexander Ilic. 2016. Understanding the impact of personality traits on mobile app adoption – Insights from a large-scale field study. *Computers in Human Behavior* 62 (Sept. 2016), 244–256. https://doi.org/10.1016/j.chb.2016.04.011
[74] Rayoung Yang, Eunice Shin, Mark W. Newman, and Mark S. Ackerman. 2015. When Fitness Trackers Don't 'Fit': End-User Difficulties in the Assessment of Personal Tracking Device Accuracy. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Osaka, Japan) *(UbiComp '15)*. Association for Computing Machinery, New York, NY, USA, 623–634. https:

//doi.org/10.1145/2750858.2804269
[75] Peifeng Yin, Ping Luo, Wang-Chien Lee, and Min Wang. 2013. App Recommendation: A Contest between Satisfaction and Temptation. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining* (Rome, Italy) *(WSDM '13)*. Association for Computing Machinery, New York, NY, USA, 395–404. https://doi.org/10.1145/2433396.2433446