Investigating In-Situ Personal Health Data Queries on Smartwatches

BRADLEY REY*, University of British Columbia, Canada
BONGSHIN LEE, Microsoft Research, USA
EUN KYOUNG CHOE, University of Maryland at College Park, USA
POURANG IRANI*, University of British Columbia, Canada

Smartwatches enable not only the continuous collection of but also ubiquitous access to personal health data. However, exploring this data in-situ on a smartwatch is often reserved for singular and generic metrics, without the capacity for further insight. To address our limited knowledge surrounding smartwatch data exploration needs, we collect and characterize desired personal health data queries from smartwatch users. We conducted a week-long study (N = 18), providing participants with an application for recording responses that contain their query and current activity related information, throughout their daily lives. From the responses, we curated a dataset of 205 natural language queries. Upon analysis, we highlight a new preemptive and proactive data insight category, an activity-based lens for data exploration, and see the desired use of a smartwatch for data exploration throughout daily life. To aid in future research and the development of smartwatch health applications, we contribute the dataset and discuss implications of our findings.

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

Additional Key Words and Phrases: Smartwatch, wearable, personal health informatics, data exploration, in-situ, diary study

ACM Reference Format:
Bradley Rey, Bongshin Lee, Eun Kyoung Choe, and Pourang Irani. 2022. Investigating In-Situ Personal Health Data Queries on Smartwatches. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 179 (December 2022), 19 pages. https://doi.org/10.1145/3569481

1 INTRODUCTION

Smartwatches have become highly capable and increasingly independent devices, equipped with a wide range of sensors, computational power, network capability, and various interaction modalities (e.g., touch and speech). Due to their wrist-worn form factor, people wear smartwatches during most of their waking hours, collecting extensive personal data in a variety of daily activities (e.g., while jogging, working, or on-the-go). Smartwatches have the potential to provide their users with immediate access to personal and situated insights drawn from the collected data. As a first step to realizing such potential, we set out to explore *personal health data queries* people look to ask directly on their smartwatch throughout daily life.

Prior works on personal health data and the mechanisms to support data exploration through visualization and interaction have largely been studied for desktop [13, 27] and smartphone [41] devices. These studies

*Authors noted conducted the majority of this work while affiliated with the University of Manitoba.

Authors' addresses: Bradley Rey, bradley.rey@ubc.ca, University of British Columbia, BC, Kelowna, Canada; Bongshin Lee, bongshin@microsoft.com, Microsoft Research, Redmond, WA, USA; Eun Kyoung Choe, choe@umd.edu, University of Maryland at College Park, College Park, MD, USA; Pourang Irani, pourang.irani@ubc.ca, University of British Columbia, BC, Kelowna, Canada.

2474-9567/2022/12-ART179 $15.00
https://doi.org/10.1145/3569481

often focus on a longer form of data exploration, which a smartwatch is not ideally suited for. Focusing on the smartwatch, research that has influence on personal health data exploration has been mainly conducted under two domains: (1) understanding the needs and preferences of smartwatch users as it pertains to their personal health data [1, 8, 35, 36, 51] and (2) data visualization [1, 5, 11, 29, 49, 52] and interaction [50] techniques to mitigate the limitations caused by the reduced smartwatch screen size. However, we still lack a broader empirical understanding of the data query needs of smartwatch users, especially within the in-situ and daily contexts in which a smartwatch is used.

Our use of smartwatches throughout daily life [56, 69] is likely to influence the type of information and exploration desired. Unlike typical data exploration on smartphones or desktop applications, smartwatches warrant unique and complementary forms of in-situ data exploration, potentially with spoken queries. Ideally, these queries could allow for new forms of accessing information, which could provide greater benefit for a smartwatch user along their personal health journey. We note, throughout this work our use of the term *data exploration* refers to this envisioned form of use for smartwatches, enabling queries that are lightweight and transient as compared with longer-form data exploration. To exemplify this, we provide a single brief scenario of a self-tracker Charles, who collects personal health data on his smartwatch. While out for a walk, Charles feels as though he is getting some great exercise, and asks his smartwatch whether he is walking faster than on his previous walk; he is, so Charles keeps up the pace and is excited to further explore his performance over the last few months later in the day. In this scenario, while Charles is walking, he briefly seeks a lightweight query of his personal health data, beyond what is currently capable on today's devices, directly on his smartwatch, which in turn can be seen to influence his current activity.

To understand, and begin to support, a broader level of in-situ data exploration on a smartwatch, our work aims to first learn the exploratory queries that smartwatch users desire throughout their daily lives (i.e., *what* data queries are meaningful to users). We additionally look to better understand any implications of being in-situ (i.e., *when/where* data exploration is conducted). Finally, we reflect on the mechanisms used by participants to submit their personal health data queries (i.e., *how* exploration can be undertaken on a smartwatch). More specifically, we focus on the following two research questions. **RQ1:** What exploratory queries do smartwatch users have of their collected personal health data, as desired throughout their daily lives and on their smartwatch? **RQ2:** How does being in-situ play a role in influencing personal health data queries of smartwatch users?

To answer these research questions, we conducted a week-long study with Apple Watch users (N = 18), and concluded with a final interview. Throughout the week, we elicited desired queries from participants that would allow them to better explore and have access to their collected personal health data on their smartwatch. Using a custom built data collection application, participants reported responses throughout their daily lives, which included a natural language query and current activity information. We learned that participants not only looked to utilize their smartwatch for data exploration temporally surrounding a tracked physical activity [1], or in-situ, but also found it beneficial to use the smartwatch throughout daily activities for immediate and often discrete exploration. Furthermore, we highlight a new Preemptive and Proactive insight category, and find expanded Current Status as well as Contextual insight categories when compared to previous works [1, 12, 13, 45]. Finally, in addition to seeking information through temporal filtering, participants often looked to explore their data through an activity-based lens.

The key contributions of this work are twofold. **C1:** An empirical study to capture queries and in-situ factors that pertain to personal health data exploration on smartwatches. We provide a dataset of the 205 natural language queries captured through our study[1]. **C2:** A qualitative and quantitative characterization of the queries sought by smartwatch users and in-situ factors that influence their personal health data exploration on their smartwatch.

[1]https://smartwatch-personal-health-data-queries.github.io/

Through the analysis of the reported responses, we discuss implications for the design of future smartwatch health applications that can provide a more unique and complementary form of personal health data exploration.

2 RELATED WORK

In this section, we cover related work in the areas of: (1) general personal health data exploration, (2) smartwatch usage behavior, (3) the smartwatch as a tool for personal health data collection and exploration, and (4) in-situ data collection methods.

2.1 Personal Health Data Exploration

Exploring personal health data to gather insight into one's personal health journey has long been a goal of quantified selfers [14]. More recently, this exploration has become common place for a more general population, largely due to the broader access to and improved performance of data collection tools and mobile health (mHealth) applications. Typical data exploration on non-wearable devices often involves longer interaction cycles, and thus can be considered as the primary action being performed. mHealth applications on devices such as smartphones, tablets, and desktops often allow for visual elements, overviews, filtering, and interaction with personal health data, enabling personal and unique insight and broader understanding of the data [2, 13, 25, 27, 41]. In contrast, smartwatches can be used within daily activities (e.g., working, walking/running, shopping, while in transport), where the activities are a person's primary focus with desired data exploration being secondary [4]. These contexts of use require a lightweight and transient form of data exploration for immediate reactionary insight [1] and quick information needs [4]. This form of smartwatch data exploration can then be seen as complementary, yet equally important to drive daily decisions and insights, to the aforementioned typical data exploration on non-wearable devices. We envision the smartwatch, taking benefit from its body-worn and readily-accessible nature, to be a device that can provide this lightweight and transient style of data exploration, anytime and anywhere. To provide increasingly appropriate and unique personal health information on a smartwatch, we aim to understand what queries people have and the contexts surrounding these queries.

2.2 Smartwatch Usage

In recent years, research has abundantly explored the usage scenarios of smartwatches and fitness trackers [28, 38, 46, 53, 56, 69, 70]. Within these works, smartwatches are generally used for time checking, notifications, and activity tracking. In addition, these types of usage are highly discrete in their nature on a smartwatch [53], and take place in many locations and during many different day-to-day activities [56, 69]. Focusing on activity tracking, as this provides us with personal health data, Chun et al. [16] found that activity tracking was the second highest smartwatch use case, with smartwatch interaction eclipsing that of a smartphone during times of activity and multi-tasking (i.e., while in-situ). Due to the on-body location, in-situ interactions with the smartwatch can happen with relative ease [38]. Furthermore, while the majority of smartwatch interactions have been found to be less than 5 seconds long [56, 69], activity and workout interactions can be longer with interactions in the 5 to 10 seconds range [42, 69] and have been found to increase up to 18 seconds on average [56].

Usage typically declines over a period of time for both activity trackers [18, 20, 30] and smartwatches [48, 53, 63]. This decline in use is mainly due to the mismatch of expectations/goals and actual offerings [18, 20, 30], inadequate support for goal setting [30], and limited functionality [48]. However, researchers have found that perceived usefulness through self quantification [15, 63], convenience characteristics of the smartwatch [53], increased *healthtology* (using technology for health conscious decisions and motivation) [24], and the active complementing of smartwatch and smartphone applications [24] provide means to motivate continued use.

In terms of smartwatch mHealth applications that can provide benefit to the user, there are many facets to consider from these aforementioned works. First, there are a broad array of smartwatch usage scenarios throughout peoples' daily lives. These interactive instances could further benefit from, and even influence, personal health data exploration. Furthermore, the transient nature of smartwatch interaction is important. Smartwatch mHealth applications should allow quick and easy access to desired information, without the need for prolonged or complex interaction which is often the case for typical data exploration. Lastly, to motivate continuous usage and increase perceived usefulness, it is important to develop innovative approaches for smartwatch personal health data exploration, which can complement their smartphone counterparts.

2.3 The Smartwatch as a Tool for Personal Health Data Exploration

Smartwatches allow for the ubiquitous collection of personal health data and activity information. People view this data through mHealth applications that take advantage of this data collection. We note that reasons for collecting such personal health data include accountability and participatory interaction [71], goal- and performance-based use, and exploration of combinations and comparisons of data [1, 8]. Furthermore, the data is often explored temporally around activities, either before, during, or immediately after an activity [1, 56]; this reflection, using a smartwatch, temporally near an activity can be increasingly intertwined with the activity itself, and thus beneficial to influence on-the-fly decisions [1, 28, 44]. Yet, it remains difficult for users to find applications that suit their needs and health goals [60], with users citing a lack of information richness and overall usefulness regarding their smartwatch and mHealth applications [54]. Neshati et al. [51] note that smartwatch users are seeking answers to their personal health data queries that are simply not available to them; these types of desired queries are exemplified as *"how am I doing so far?"* or *"how am I doing compared to my friend, Jane?"* [45], and are not currently supported within a time frame that supports common smartwatch use. These missing features, and limited knowledge of concretely reported queries, hamper full engagement and a broader range of benefits [30].

While a growing body of research revolves around smartwatches and other wearable activity trackers, only about 10% of such research is centered around the collected data and the ability to convey appropriate meaning and function [62]. These works often highlight the complex and visually rich personal health data represented on screen having limited usability, customization, and interaction [16]. As such, individual components of smartwatch data visualization [1, 5, 11, 29, 49, 52] and interaction [50] have been explored. However, while beneficial to their respective aspects, these works are either specific to a single exploratory purpose or chart type, provide limited unique-to-user insight, or look to replicate current data exploration tasks and needs. To gain an understanding of the broader needs of smartwatch users pertaining to their personal health data this work captures queries, desired to be answered on the smartwatch, from participants throughout their daily lives. These queries can be seen as a direct means for further accessing and exploring one's personal health data. This knowledge will allow us to identify design opportunities for smartwatch mHealth applications that can play a complementary and important role in personal health data exploration.

2.4 In-Situ Data Collection Methods

Typical in-lab data collection methods, such as interviews, surveys, and focus groups are often subject to recall bias [32]. Thus, ecologically valid methods of data collection have been utilized in research across domains. These methods include, Diary study [6, 9, 17, 23, 31, 39, 64] and Experience Sampling Method (ESM) (or the equivalent Ecological Momentary Assessment) [19, 22, 43, 68]. Both methods often utilize a pre-built survey or questionnaire for participants to respond with, but a notable difference between Diary and ESM studies is the nature in which data is meant to be captured. ESM studies utilize notifications sent to participants, where data capture is intended to be done in immediate succession to the alert. On the other hand, Diary studies allow for the collection of in-situ self-report data whenever participants have a response to report. Notifications are then used as reminder rather than a trigger to submit responses. To further reduce data capture burdens, researchers have begun to utilize smartphones and smartwatches for successful data collection [33, 40, 72]. The always-available and body-worn nature of a smartwatch allows for easy access to the data collection tool. Furthermore, notifications on smartwatches have a high level of awareness from the participant [10, 56], have shown to allow for higher response rates, and are perceived as being less distracting during daily life [34, 58].

3 USER STUDY

The goal of this study was to collect and better understand personal health data queries smartwatch users have throughout their daily lives. Our in-situ data collection spanned the course of a week, in which we employed a diary study method. As part of the study, we installed a custom data collection application on participant's smartwatches, through which they entered a desired personal health data query using natural language (either spoken or written), as well as additional activity information. Participants provided responses throughout their daily lives, specifically when they felt it would be beneficial to access such information concerning their personal health data directly on their smartwatch. Ethics for this study was obtained from and approved by our institutional review board.

3.1 Participants

We recruited 18 participants (P1–P18; ten females and eight males) from Reddit. We advertised our study across subreddits relating to personal health as well as a number of general city subreddits across Canada. Our inclusion criteria were those who (1) were aged 18 years or older; (2) own an Apple Watch Series 3 or newer with watchOS 7 or higher installed, and have a paired iPhone; (3) have access to Zoom and a stable internet connection; (4) are native English speakers; (5) have no motor, visual, or speech impairments; (6) currently collect at least one of the following personal health data: sleep, nutrition, physical activity, steps, calories, women's health, and/or mindfulness data on their Apple Watch; and (7) have been regularly collecting personal health data on their Apple Watch for at least three months.

In appreciation for their time and effort, participants were offered up to ~$30 USD in local currency. The amount a participant received was not tied to the number of responses, but rather the number of participation days. We provided the equivalent of $7.50 USD for attending the introductory session, and another equivalent of $7.50 USD for attending the final interview. During the week long data collection, we added an additional equivalent of $2.15 USD for each day for which a participant provided at least one response. We provided compensation after the final interview, or upon withdrawal, in the form of an electronic Amazon gift card.

3.2 Data Collection Method

To understand what personal health data queries lay users have, researchers have previously used focus groups, lab based experiments, and questionnaires [1, 59]. While these methods provide a positive general understanding, we aim to mitigate the potential for recall bias through concretely reported, in-the-wild, responses. Within our work, desired responses could arise at any time within a participant's daily life. Therefore, we utilized a Diary study method combined with experience sampling method (ESM)'s random-interval sampling. We note, both fixed interval-based and event-triggered collection methods would have restricted responses to specific and predictable times or to certain activities respectively. A Diary study method combined with ESM's random-interval sampling allows for a wide range of boundary pushing and ecologically valid queries to be captured, during a range of times and activities, without restriction, as they are deemed beneficial to the participant.

3.3 Data Collection Application

We created a data collection application[2] for the Apple Watch using Swift, and deployed it via Apple's TestFlight beta program to each participant's individual Apple Watch. We chose only the Apple Watch due to the immensely

[2]https://github.com/reyb/Personal-Health-Query-Recorder

simple logistics in installing our application remotely. This also ensured consistency in data collection procedures. Our application utilized Google's Firebase Realtime Database to collect and store the responses submitted. As our goal is to capture queries for personal health data exploration on smartwatches, the application was designed purely for data collection; it did not answer the desired queries from our participants.

3.3.1 Data Collection Questions. Our application consists of up to four questions for participants to respond to. All questions were required to be answered when submitting a response. The questions and input methods were designed to support fast and easy reporting of responses while mitigating interaction difficulties on a small-screen device. This included the use of natural language reporting for open-ended questions (Q1 and Q2), leveraging the flexibility and ease in reporting ability [40], and single screen options for Q3 and Q3-1. Our application captures the following information, illustrated in a flow diagram shown in Figure 1 and described below:

Q1 (open-ended): What question or command do you have of your health data? This allowed us to capture the personal health data query from the participant. Q1 elicited a query regarding what participants were interested in quickly exploring or accessing on their smartwatch. Participants could either speak or type a query using the Apple Watch's built in text-entry methods. Upon recording the query, it could be reviewed on-screen in real-time and repeated to correct errors, if needed.

Q2 (open-ended): What are you currently doing? This helped us gain general contextual and daily activity information of the participant's daily life at the time of recording a desired query. For simplicity on the part of the participant, we allowed for natural language input in the same manner as the first question.

Q3 (dichotomous): Is your response related to your current activity? As a personal health query may or may not relate to the current activity being performed, this question allowed us to better understand the association between a participant's current activity and their personal health data query given in Q1. Either a "Yes" or "No" answer could be selected.

Q3-1 (multiple choice): Where are you in your activity? This would appear to the participant only if they selected "Yes" in the previous question. From this question we aimed to gain further knowledge, understanding the in-situ moment surrounding an activity, that a need for exploring personal health data arises. "Before," "During," or "After" could be selected. In order to better understand when a desired query may be temporally related to an activity, we ensured participants understood that our definition of relation could also include just before starting and after completion of an activity. For example, a participant could be going to the gym looking to soon start their workout (Before), actively engaged in their workout (During), or heading back to the change room having just finished (After). By further exploring this time distinction, we can introduce a greater level of granularity and aim to understand when to provide exploratory capability or further insight to smartwatch users.

3.3.2 Reminders. To elicit many responses, our application employed two forms of reminders. First, we used push notifications. We customized these notifications to each participant, based on their self-declared wake and sleep times. Furthermore, the notifications were systematically random in that a notification would be sent at a participant's declared wake time, and then concurrently sent throughout the day between one to two hour intervals (chosen at random) after the prior notification. Notifications in this manner were repeated until the participant's declared sleep time. This method was chosen to ping participants at different times throughout each day, ideally attempting to remind them of the study at different in-situ moments within their daily lives. Second, our application also leveraged a watch-face widget, as seen in Figure 1 (right). The small circular widget was continuously displayed on a participant's home watch-face and additionally provided a counter of the number of responses a participant had submitted throughout the study.

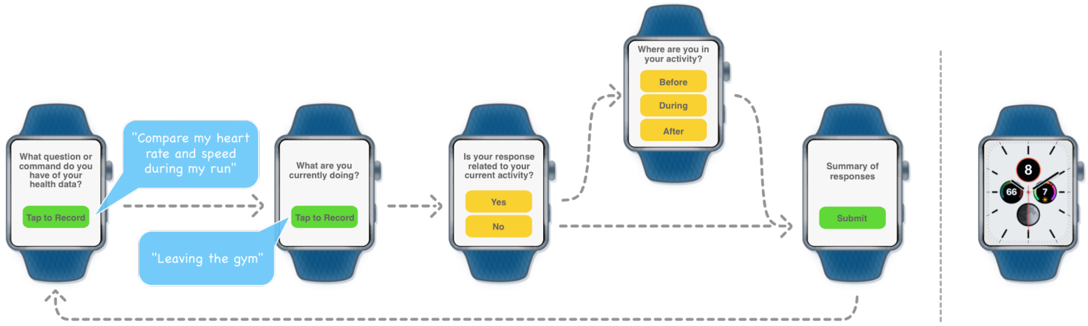


Fig. 1. (Left) Flow diagram of the questions asked within our data collection application. (Right) Watch face widget in the top center. We used a small circular widget which could be placed on a range of watch face styles, providing a counter of the number of responses a participant had submitted throughout the study.

3.4 Study Procedure

Our study included three stages: an introduction and tutorial session, a seven-day in-the-wild data collection, and a final interview. The procedure and study materials were iterated upon during two pilots with people who were recruited in the same means as our participants, thus meeting our study's inclusion criteria. Participants provided consent at the start of the study.

3.4.1 Introduction and Tutorial Session. To start the study, the participant joined a Zoom meeting, ~45 minutes long, where we introduced and acknowledged their interest and participation in our study. Participants were encouraged and asked to interject with any comments and/or questions during the meeting. The researcher shared presentation slides (please refer to the supplemental material) via the screen sharing functionality. The goal of the project was expressed to the participant along with other important remarks before they completed a demographic survey. Then, the researcher guided the participant through the installation and setup process of the application on the participant's own Apple Watch to be used throughout the data collection stage of the study. This included setting up notifications and the watch-face widget for reminders. The researcher then gave participants a walk-through of the application through an on-screen emulator running the application on the researcher's computer and shared via screen sharing. Upon completion of the walk-through, the researcher gave participants unrelated-to-the-study practice trials to ensure the application worked smoothly and any technical issues related to data collection were appropriately handled.

Finally, the researcher gave an explanation on the potential queries and purpose of the study. The researcher asked participants to provide queries that they deem as beneficial, without worrying about current technological limitations. Queries would ideally allow participants to better explore or access their personal health data directly on their smartwatch. The researcher also asked participants to only provide a query when it both arose within their daily lives and was deemed to be a data exploration task they would like to perform on their smartwatch. The researcher explained to participants that feedback or answers to their queries would not currently be given, however they should envision receiving this directly on the smartwatch. No specific examples were provided to the participants, so as not to bias their potential queries. However, high-level categories of health data exploration (e.g., history of data, goals/performance data) were discussed to invigorate ideation.

3.4.2 Data Collection. Participants used the application to submit responses over the course of the next seven days, and were instructed to wear their watch as they normally would. A response could be submitted at any time throughout the day. We did not require a minimum number of responses throughout the study, to not elicit forced responses from participants. Due to the potential for participants to be active when a query arose, we instructed them to only provide a response through the application when it was safe for them to do so. Notifications sent each day acted solely as reminder of the study.

3.4.3 Final Interview. After the final day of data collection, a Zoom meeting was held where a researcher conducted a semi-structured interview with each participant. The meeting was audio recorded and later transcribed. The goals of the interview were: (1) to gain a better understanding of the smartwatch's role within each participant's health journey; (2) to explore additional information regarding in-situ smartwatch use for personal health data exploration; and (3) to discuss preferences in terms of interaction and visualization when exploring personal health data on a smartwatch. To aid in recollection, a report of each participant's queries were shown to them via Zoom's screen sharing functionality. Finally, the researcher answered any remaining questions from a participant, thanked them, and provided the compensation. Each interview took between 30 and 45 minutes.

4 ANALYSIS AND RESULTS

4.1 Analysis

A total of 229 responses were logged through our application. First, we manually inspected the data, curating a dataset that only included valid responses. Through this process, we discarded 24 responses which fell under three categories: (1) the query had no specific element of collected personal health data (e.g., *"Am I over the food poisoning from yesterday?"*); (2) the query was related to smartwatch functionality rather than personal health data exploration (e.g., *"Is there a better way to track active minutes?", "How much battery does tracking a walk use?"*, *"Is there a way to account for temperature while working out?"*); (3) the query was action based and did not allow for exploration (e.g., *"Set my bedtime for 12 AM and wake me up by 8 AM"*, *"Start outdoor run"*, *"Record 96 ounces of water for the day"*). After this process, 205 valid responses remained for analysis.

We qualitatively analyzed the valid responses, containing a query, activity of the participant, and relation to this activity through an open coding process. To do this, we followed the same approach as used by Srinivasan et al. [66]. To specify, two researchers first explored the reported queries and activities for broad themes, subsequently creating a coding schema. Then, the same researchers individually coded random subsets of the data after which they came together to compare results for agreement. They refined the schema and codes, and individually coded a new random subset of data, until an 85% agreement was reached. Once the researchers came to agreement, the data was independently coded in full using the mutually-agreed upon codes, again working together to reach full consensus as needed.

4.2 Results

Table 1 summarizes the demographic, smartwatch usage, and health data collection information, as well as response counts of our study participants. Participants were aged from 18 to 56 (M = 29.8) and held a range of occupations. At the time of conducting the study, participants had collected personal health data for an average of 39.3 months (SD = 32.7 months) and had used a smartwatch for an average of 31.3 months (SD = 26.0 months). Collected personal health data by participants mainly focused in two categories, physical activity (18/18 participants) and sleep (7/18 participants) data. Of our 205 valid responses there was an average of 11.4 responses per participant (SD = 5.5; Min = 3; Max = 22).

We coded our valid responses, leading to the following five dimensions: (1) personal health data insight category, (2) current daily activity, (3) whether the query was related to the activity, (4) the time in activity if related, and (5) the query type. Taken together, these can enable a better understanding of the characteristics surrounding desired smartwatch data queries, combined with the types of insight that are of interest during certain daily activities. Below we detail each dimension and their codes.

Table 1. Summary of demographic information, health data collection, smartwatch usage experience, and number of responses reported from our study participants.

Alias	Age	Gender	Occupation	Health Data Collection	Smartwatch Usage	Collected Data	# of Study Responses
P1	19	F	Customer Service	1y 2m	1y 2m	Sleep, Physical Activity, Women's Health	8
P2	31	F	Care Coordinator/Dementia Counsellor	10y 0m	2y 6m	Sleep, Physical Activity, Women's Health	10
P3	38	M	Communications Advisor	5y 6m	5y 6m	Sleep, Physical Activity	16
P4	35	F	Teacher	3y 0m	3y 0m	Physical Activity, Women's Health	9
P5	33	M	Information Security Specialist	5y 1m	1y 0m	Sleep, Physical Activity	9
P6	24	F	Student	1y 0m	0y 8m	Nutrition, Physical Activity	3
P7	42	M	Claim Evaluator	3y 0m	3y 0m	Physical Activity	15
P8	25	F	Scientific Evaluator	0y 3m	0y 3m	Physical Activity	6
P9	36	F	Post Doctoral Fellow	0y 8m	0y 8m	Physical Activity, Women's Health, Mindfullness	20
P10	30	M	Student	7y 0m	7y 0m	Physical Activity	12
P11	23	F	Student	2y 3m	2y 3m	Physical Activity, Mindfullness	16
P12	56	F	Retired Lawyer	1y 9m	1y 9m	Physical Activity	7
P13	22	F	Educational Assistant	6y 4m	6y 4m	Physical Activity, Mindfullness	18
P14	21	M	Student	2y 1m	2y 1m	Physical Activity	22
P15	18	F	Customer Service	0y 5m	0y 5m	Nutrition, Physical Activity, Mindfullness	11
P16	25	M	Software Engineer	6y 7m	6y 7m	Sleep, Nutrition, Physical Activity, Mindfullness	5
P17	40	M	City Planner	1y 11m	1y 11m	Sleep, Physical Activity, Mindfullness	15
P18	18	M	Student	0y 11m	0y 11m	Sleep, Physical Activity	3

4.2.1 ***Expanded Personal Health Data Insight Categories.*** Through the queries desired by participants, we coded the overarching insight category for which each query aligned. Table 2 shows the codes, descriptions, and selected queries. We expanded upon categories from previous works [1, 12, 13, 45] to assign our collected queries into the following codes: Current Status or Value (89, 42.9%), Historical or Trend (67, 32.7%), Combination or Comparison (59, 28.8%), Goals or Performance (57, 27.8%), Preemptive and Proactive (47, 22.9%), and Contextual (24, 11.7%). These codes are not mutually exclusive, and thus a query can have multiple codes (and as such the percentages above are individually calculated from the 205 total queries). For example, *"How much dancing do I need to do to burn 800 calories?" (P18)* can be seen as fitting into both the Preemptive and Proactive and the Goal and Performance categories. Of the six codes, participants reported queries in a minimum of 3 codes and maximum of 6 (M = 5.0, SD = 1.0).

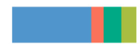
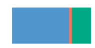
Current Status or Value insight was often about more than just the simple metrics captured. Amini et al. [1] found that participants explored step count, distance, calories, pace, speed, and heart rate during an activity. While this remains to be seen in our study, we found a broader definition of current metrics desired; these included heart rate zones, total values from activities throughout the day, peak values or fluctuations of metrics throughout the activity, and aggregated values such as perceived exertion (see Table 2 for specific queries). These examples, which can be seen as increasingly unique-to-user, should be considered within smartwatch health applications to expand the usefulness and benefit to a broader range of smartwatch users.

Table 2. Breakdown of the query categories which provide insight to users. We note that these categories are not mutually exclusive, thus an example given could fit into multiple categories. ** denotes a new insight category found in our work, while * denotes an expanded insight category as compared to [1, 12, 13, 45]. Q is a question; C is a command.

Response Category (# of Responses, % of Total, # of Participants)	**Description**	**Example Responses [all directly quoted]**
Current Status or Value * (89, 42.9%, 17)	Current, single value, metric that is collected and/or aggregated on the smartwatch to be given to the user.	(Q) Am I over or under my calorie goal at the moment? (P15)* (Q) How many calories did I burn that workout? (P4) (Q) How many steps am I at today? (P4) (Q) What was my peak heart rate during my workout? (P15)* (C) Give me a report for my readiness for activity. (P16)
Historical or Trend (67, 32.7%, 18)	Previously collected metrics, prior to the current day's or activity's. Can often be used to explore changes over time.	(Q) How long on average does it take me to fall asleep? (P3) (Q) How many steps have I taken this week? (P9) (C) I would like to check a trend in my sleep in the past seven days. (P1) (C) Show me my body weight trends for this month. (P6)
Combination or Comparison (59, 28.8%, 17)	Combine and/or compare two or more different measured values. These can be done over time, between metrics, between activities, or between oneself and others.	(C) Compare my running stats with the same time last year. (P17) (C) Show me a graph of my runs both time and distance in 2021. (P11) (Q) Does my walking pace change when I walk with someone else? (P12) (Q) How many calories were burned in today's work out compared to yesterday? (P6)
Goals or Performance (57, 27.8%, 14)	Goals, such as for steps, calorie intake, calories burned, distance travelled, etc. Performance stems from completing a goal, as well as quality metrics such as fast/slow or best/worst.	(Q) How fast did I finish my 1st kilometer of my hike today? (P6) (Q) What was my best kilometer during my run? (P4) (Q) Which activity had the highest calories burned per minute? (P18) (Q) How many kilometers do I need to walk to get 10,000 steps? (P4) (C) Tell me when I reach a nine minute walking pace. (P12)
Preemptive and Proactive ** (47, 22.9%, 15)	Advice or information that will allow one to be preemmptive or proactive in making decisions and/or to prepare for an event in the future.	(Q) How long do I need to run three times per week to achieve the November challenge? (P17) (Q) Based on my current biometrics when will be the best time for me to work out today? (P10) (C) Give me a suggested work out based on my readiness score. (P16) (Q) Is there a day of the week I am more likely to beat (friend) in our fitness challenge? (P12) (Q) How much dancing do I need to do to burn 800 calories? (P18)
Contextual * (24, 11.7%, 9)	Impact and/or affect of an external or collected metric on another.	(Q) During which song did I burn the most calories? (P18)* (Q) How is the air quality affecting my walk? (P16) (Q) Is my cycle affecting my sleep? (P2) (Q) Is my running pace slower in the days following a strength training workout? (P13)* (Q) What is the impact of my sleep in my running? (P17)*

Contextual exploration was utilized by participants to find cause and effect between a range of data. Choe et al. [13] discussed participants' interest to include and explore the effect that external data such as time of day, location, or weather can have on their own collected personal health data. While our captured queries from participants garnered similar contextual information, we also note that participants were inclined to look for cause and effect using their own collected data as context. For example, *"Is my cycle affecting my sleep?" (P2)*, *"Does my walking pace change when I walk with someone else?" (P12)*, *"What is the impact of my sleep in my running?" (P17)*, and *"Does weightlifting focusing on different muscle groups affect my heart rate?" (P14).* As participants used

Table 3. Summary of the daily activities participants were undertaking at the time of a response, and the relation of the reported queries to these activities: **not related**, **before**, **during**, **after**.

Daily Activity	Description & Examples	# of Responses (% of Total), # of Participants	Query Related to Daily Activity
Physical Activity	Body movement that requires more energy than resting. (e.g., dancing, walking, running, weight training, yoga, sports, etc.)	70 (34.1%), 16	
Self Care	Activities that pertain to normal day-to-day human function. (e.g., cooking, eating, chores, morning ready routine, etc.)	44 (21.5%), 12	
Work	Fulfilling duties either for job or school. (e.g., job based tasks, studying, attending meetings, etc.)	29 (14.1%), 12	
Leisure	Activities performed for relaxation and fun. (e.g., reading, watching TV, lounging, etc.)	25 (12.2%), 14	
Sleep	The absence of wakefullness. (e.g., napping or nightly rest)	17 (8.3%), 9	
Transportation	Moving from point A to point B. (e.g., driving, taking the bust/metro, taxi, etc.)	10 (4.9%), 5	
Other	Activities that do not fit within the prior categories. (e.g., within the study interview, due to mHealth app notifications)	10 (4.9%), 5	

the smartwatch to capture a range of health and activity data either automatically or through a discrete input, they desired to explore context surrounding these captured instances on the smartwatch.

Preemptive and Proactive queries, a new form of insight brought forward from our study compared to previous works [1, 12, 13, 45], make up ~20% of our collected data from 15 of 18 participants (see Table 2 for specific queries). Smartwatches can provide influential decision making through captured metrics, temporally related to activities [1, 45]; this is often noted to be from simple viewing of metrics displayed on screen (e.g., a runner glancing at their smartwatch noticing their pace is not currently at their desired value can up their speed to improve). Our participants, however, were looking for a wider range of influential exploration from their smartwatch, such as to help choose a workout for the day, plan an activity based on goals, or to pick up on elements that they alone may not be able to predict. This form of insight was seen as a means for preparing oneself for a future event rather than simply reflecting on current or past metrics. Participants were looking to utilize these Preemptive and Proactive insights from their smartwatch to influence immediate and in-situ decision making, as well as for some daily and longer term planning. This longer term planning was often seen as tied to the Goals or Performance category, as 23/47 Preemptive and Proactive queries were also to gain insight on how to achieve immediate or future goals. Finally, while about a half of the queries in this category were temporally related to an activity, we further note that Preemptive and Proactive queries were desired on the smartwatch throughout the day.

4.2.2 ***Smartwatch Personal Health Data Queries Desired Throughout Daily Activities.*** Analyzing the daily activities of participants led to seven codes. These codes categorize the current activity being performed in a participant's daily life, during which their desire for exploration on the smartwatch arose. These codes include: Physical Activity, Self Care, Work, Leisure, Sleep, Transportation, and Other; the codes, their descriptions, and counts can be seen in Table 3. Codes were mutually exclusive, and thus each response was given a single code. Of the seven codes, participants reported responses in a minimum of 2 codes and maximum of 6 (M = 4.1, SD = 1.4).

We asked participants whether their personal health data query was related to their current activity as reported above (Q3), and thus in-situ. This was either a Yes (107, 52.2%) or No (98, 47.8%) answer; examples from our captured data include *"Leaving the gym" - "Show me my heart rate chart from today's gym session"* as being related while *"At work" - "What was my fastest kilometer in my run?"* as not. We see an almost equal distribution overall, however, when combined with the activity we see a distinction; the most drastic being in relation to Physical Activity. Here, 97% of queries reported were related to the Physical Activity being done. In contrast, all other daily activities, aside from Sleep, provided time within daily life for increased insight and reflection that was unrelated to the participant's current activity. Within our captured responses, 16 of 18 participants found it beneficial to report a query both related and unrelated to the current activity they were doing.

When queries were related to the daily activity being done by the participant, we additionally asked whether they were just Before (24, 22.4%), During (38, 35.6%), or After (45, 42.0%) the activity. The results collected follow closely with our demographic survey which asked when participants aim to explore their data on their smartwatch, Before (4/18 participants), During (10/18 participants), and After (15/18 participants). Participants noted during the interviews that exploration Before activities could only do so much to affect the activity collecting the health data once started. Thus, in relation to our insight categories, broader Preemptive and Proactive as well as general Goal insight was most often queried Before an activity. During an activity, specifically for Physical Activity, expanded insight beyond simple metrics pulled participants away from being in the moment and focused on the activity at hand, thus was not as desired. P5 mentioned *"It wasn't much [exploration] like during the workout 'cause in the workout I found that is more like just concentrating on whatever I was doing and I didn't really have any questions to ask."* This is reflected in the Current Status or Value insight category being the most queried During an activity. In fact, of the queries that took place surrounding Physical Activity, 26.5% were During the activity, or in-situ, of which 68% of these were simply to understand a Current Status or Value. Finally, exploration after an activity allowed for immediate reflection to take place which could help influence future activities. Interestingly, participants most often looked to perform data exploration surrounding Physical Activity after the completion of and regarding the activity itself (32/66 queries). This reflection After an activity was mainly of the Current Status or Value, Historical or Trend, and Combination or Comparison insight categories.

However, from the interviews held with participants, we note that these captured results may not provide the entire picture. First, while our results provide a general understanding of what queries and when these queries are desired, we note that across combined insight categories, daily activity, and relation to the activity, our results show queries reported in 128 unique combinations of these. This highlights the deeply unique and personal aspect to personal health data exploration needs. Second, two participants suggested that while they may utilize different in-situ moments within an activity to explore their data, the activity itself was not always the determining factor when they aim to explore their data on the smartwatch. Notably, current overarching goals set by a participant and how long they had used a smartwatch for tracking data could affect the type and time of their exploration. As P7 discussed, *"Because I've been using it [the smartwatch] for like a couple years, uhm, I think I'm pretty good at like knowing what kind of workout will make me hit my calorie burn goal or get the steps I need or those kinds of things."*

4.2.3 ***Exploration Through Commands; Information Immediacy Through Questions.*** Guided by the definitions and codes created by Srinivasan et al. [65] for natural language data exploration, we found that participants' queries were framed either as a Question (173, 84.4%) or Command (32,[3] 15.6%).

The majority of Commands, provided by over a half of participants (11/18), were of the Historical or Trend as well as Combination or Comparison insight categories (21/30). Examples include *"Show me a graph comparing my caloric intake over the last week" (P7)*, *"Show me a graph of my runs both time and distance in 2021" (P11)*, *"Give me a report for my readiness for activity" (P16)*, and *"Compare cycle data from today to the same day in my last*

[3]We categorized two instances of queries from P18, "Summary of my sleep cycles," as an (implicit) Command.

cycle" (P2). Through these Commands, we can see that the desired outcome of the participant is not explicitly clear (i.e., while we can try to provide an optimal visualization, there is not a discrete answer that can be given). Additionally, these Commands often suggested that participants had the intention to further explore or view a larger range of collected health data, often through on-screen visual representation directly on the smartwatch.

Conversely, a Question was often much more direct and closed-ended, with the intended result of the insight seemingly known to the participant. All participants provided queries in the form of a Question. Examples include *"Have I stood up this hour?" (P17)*, *"How long on average does it take me to fall asleep?" (P3)*, *"How many hours did I sit yesterday?" (P9)*, *"How many steps did I get during that 2 kilometer walk?" (P4)*, and *"What was my calories burned in the last 30 minutes?" (P3)*. These Questions benefit the type and length of interaction that is typically undertaken by smartwatch users [56, 69], as they can allow for direct feedback. P14 discussed during the interview, *"I think the question sort of implies immediacy [...] and I think it's the immediacy that the watch would be nice if it covered."* This finding provides us with valuable information surrounding the intent and perceived use of the smartwatch for personal health data exploration, especially while in-situ. Often times, throughout one's day, discrete and immediate insight is valuable. This insight, while potentially leading to further exploration, does not require it, and thus the use of a smartphone or desktop application is not immediately needed.

4.2.4 ***Activity Based Lens Used For Smartwatch Data Exploration***. The smartwatch is inherently a device that captures instances of activity (e.g., tracking nightly sleep, individual runs or walks, when food/water is consumed, when weighing yourself using a connected scale). While this can be done either automatically or through discrete input, it was seen as natural for participants to incorporate activities within their proposed queries and thus desired exploration. This was done not only for Contextual insight, but also for data filtering.

Filtering was often utilized throughout the queries recorded from participants. The first type of filtering was time based, a common and beneficial filtering technique for personal health data [41] as it is a primary dimension of the data. The second type was activity based; utilized by our participants, activity based filtering was seen in over 20% of queries; examples include *"Has my heart rate during walking decreased since I started tracking walks?" (P13)*, *"Is my running pace slower in the days following a strength training workout?" (P13)*, *"Show me a graph of my average walking speed during outdoor walks" (P3)*, and *"On average was my heart rate lower during today's run compared to yesterday's?" (P6)*. These queries allow filtering to be accomplished naturally by referencing when the smartwatch was used to track data, without the need for a user to remember or further explore specific times an event or activity occurred.

5 DISCUSSION AND FUTURE WORK

5.1 In-Situ and Non-In-Situ Preparation-for-Action

Reflecting on personal data using a smartwatch can occur increasingly close to the action, for which the reflection is related, to benefit on-the-fly decisions [1, 28, 44]. In fact, when reflection and action are related, Ploderer et al. [57] suggest there exists reflection-in-action (i.e., real time and Current Status or Value insight) and reflection-on-action (i.e., aggregation of data such as Historical or Trend, or Contextual insight) enabling both maintenance and discovery, respectively [45]. Rather than reflection-on-action only taking place in-situ and immediately after an activity, our results also showcase a need for discovery and reflection while away from an action for which the query would be related. We further postulate that smartwatch users are looking for additional discovery through preparation-for-action. This form of exploration was deemed as beneficial by our participants most often for Preemptive and Proactive as well as goal and performance insight. This preparation-for-action not only happens in-situ, immediately prior to an action, but also away from the action for which the reflection was related. Many research works focus on the in-situ exploratory capabilities of the smartwatch [1, 37, 40, 61], however expanding on the smartwatch's capabilities during non-in-situ usage scenarios could be critical for further adoption, continued use, and a range of benefits.

5.2 Query Insight Category Dependent on In-Situ Activity

While coding our data, we recognized that identical queries could lead to different insight categories. For these, the insight category of the query was highly dependent on the daily activity being performed and its relation to the query (i.e., whether the query was being reported in-situ or not). As an example from our dataset, *"What is my average walking pace?" (P3)* can imply and elicit different meaning depending on when it is asked. For instance, if the query was asked in-situ while during the middle of a walk, the answer could likely be seen as the average walking pace of only the current walk (Current Status). However, if the query was asked while sitting down at work the answer may require the calculation of the average walking pace across all walks recorded (Historical or Trend). Thus, utilizing current information available from the smartwatch's sensors as well as user-initiated activities can at times become a key component in understanding a lay person's personal health data query and information needs on a smartwatch. This can then be crucial in regards to formulating appropriate responses to allow for lightweight and transient data exploration that is beneficial to smartwatch users.

5.3 Enabling Preemptive and Proactive Insight on Smartwatches

Preemptive and Proactive exploration was a new insight category found within our collected data. We believe that this was observed due to our study setup; we emphasized to our participants that they should not worry about the capabilities of existing technologies, whereas previous data exploration work has been conducted with a working prototype despite its limited functionality [13, 27, 41]. The queries within this category were aimed towards people utilizing their personal health data to prepare for future events through system recommendations. Currently, mHealth applications and smartwatch operating systems provide a primitive form of this type of insight. Notifications, suggestions, or motivational reminders are often utilized to proactively encourage people to stand up, breathe, or move after sedentary periods. However, these are limited in their expressiveness and ability, and do not allow for data exploration to be included.

In the pursuit of enabling this form of insight, many questions arise; to showcase these, we highlight two queries from our data: *"How much dancing do I need to do to burn 800 calories?" (P12) and "Give me a suggested workout based on my readiness score?" (P16).* First, how should the answer be calculated? Calculating the appropriate answer can be a challenge, especially when little prior data is available or standards are unknown. Second, when providing an answer, how can we convey uncertainty and variance to the user? These queries often do not have a discrete answer available, with more factors and external data needing to be considered. Finally, is this form of insight ethically possible? With the ability to recommend, and ultimately have a person act upon an answer, this carries with it the importance of not misleading a person which could have ramifications. These areas of future study are important in enabling this form of beneficial and desired insight on a smartwatch, and even for a broader set of devices.

5.4 Interacting with Personal Health Data on Smartwatches

As seen from our collected data, the in-situ information needs of smartwatch users are more varied and broad than currently available. This highlights the unique nature of in-situ and day-to-day smartwatch data exploration, and brings into question how this range of exploration can be undertaken in a manner appropriate for smartwatches. Without the screen real estate for menus, toggles, scales, etc., touch interactions can be difficult on a smartwatch [26]. Therefore, we suggest utilizing a multi-modal approach, which incorporates speech and touch. This has already proven to be liked and effective for smartphone personal health data exploration [41]. Touch is a natural and common input modality, already being used for smartwatch data exploration [50], while speech allows for the conveyance of more complex queries, a low barrier in expressing intent, and flexibility in phrasing and querying [3, 21, 66, 67]. Through the use of these input modalities benefits of both could be utilized to enhance smartwatch data exploration, yet remains to be explored.

While research has focused on visual representation [1, 5, 11, 29, 49, 52] and interaction [50] of data for small screen displays, we suggest that future research should explore the potential for natural language based feedback of personal health data. This form of response is both natural and quick, ideal for the typical length and in-situ nature of smartwatch usage [56, 69] and many of the queries collected from our study. Furthermore, research has shown that auditory feedback can benefit small devices and in times when the visual system is overloaded [7, 55], much like the use of a smartwatch throughout daily life and in certain activities. Research regarding how to phrase answers, the amount of information that can be provided, and the level of insight and supplementary information that can be derived compared with and when supplemented by visualizations, should be studied.

We encourage the reader to view and use our contributed publicly available data (included in the supplemental material) in pursuit of this interaction goal. Understanding semantics used in regard to personal health data exploration and creating tools to process these queries [59], as well as determining how to convey results could lead towards a greater benefit for the smartwatch end user.

5.5 Limitations

Within our study, we recruited smartwatch users who had been collecting personal health data for a minimum of three months. While other work looking at personal health data exploration has used similar inclusion criteria [41], we do note that this does limit our findings to individuals who have at least some knowledge of their desired data exploration. New smartwatch users may prove to have different exploratory needs. We also recognize that our strict inclusion criteria resulted in the exclusion of individuals with impairments. Thus, our dataset is not fully representative of all who utilize a smartwatch for personal health data collection, exploration, and health monitoring. While our work largely provides a general understanding of personal health data queries on smartwatches, we suggest these aforementioned user groups should be studied in their own regard.

Due to the early nature of our work, we chose to not provide feedback to our participants (i.e., the answers and data representations in response to queries). This allowed us to capture a broad range of desired smartwatch data queries without influencing further potential responses. However, this design does limit us in understanding any forms of continual and serendipitous exploration on the smartwatch, such as asking a single followup question based on feedback given. During typical data exploration, one explicit query is often not enough to fully represent what is interesting to a person. Feedback given in response to a query can often lead to subsequent and unanticipated queries [47]. Moreover, this does not allow for discovery of new and interesting information beyond what is asked. While the smartwatch may not allow for lengthy data exploration, the limit to which the smartwatch can enable follow-ups, the relationship between the smartwatch and other data exploratory tools (i.e., smartphone, tablet, desktop), and the impact that given feedback has on smartwatch personal health data exploration should be further studied.

Finally, as with any elicitation study there are additional limitations to note. First, we are likely not able to capture all usage scenarios participants may experience, and for which a desire for smartwatch data exploration could arise. Examples of such times include a person preparing for a marathon or someone on vacation. Second, participants often do not realize a device's full potential limiting the range of responses submitted. For example, blood glucose monitoring has a potential future within smartwatches, yet was not a component in any queries reported by our participants. We aimed to mitigate both of these limitations through recruiting eighteen participants, each running the study for one full week (including weekdays and a weekend), and expressed for participants to not worry about current technological limitations instead focusing on queries that they desired to be possible. As such, we believe that our results remain to provide a broad understanding and new insight into queries desired, which in future can be translated to a range of usage scenarios and newly captured data.

6 CONCLUSION

This work provides an empirical understanding of queries smartwatch users have to better explore and gain insight into their collected personal health data. Specifically, through an in-situ diary study with 18 participants over a week long period, we captured queries which would allow for exploration on a smartwatch throughout one's daily life. Through the results of this data collection and a final interview with participants, we report on query insight categories, query relation to daily activities, and methods participants use for framing queries. Participants reported a desire to utilize the smartwatch for momentary and immediate personal health data exploration, not only during in-situ moments but also across a range of daily activities. We suggest several key implications for the design of smartwatch mHealth applications; including supporting preemptive and proactive exploration, expanding upon current status and contextual exploration, and allowing for activity-based data filtering. We strongly believe that the smartwatch can become a powerful tool in regards to personal health, especially as it pertains to quick and in-situ provision of data-driven insight to users. We hope that this research, along with our publicly available collected data, fosters and inspires future work in building a new generation of smartwatch applications, providing a richer set of personal health data exploration anytime and anywhere.

ACKNOWLEDGMENTS

We would like to thank all of our study participants for their time and effort involved in participation. Bradley Rey was in part supported by the Visual and Automated Disease Analytics Program. Finally, we appreciate the support during the coding process from Charles-Olivier Dufresne-Camaro.

REFERENCES

[1] Fereshteh Amini, Khalad Hasan, Andrea Bunt, and Pourang Irani. 2017. Data representations for in-situ exploration of health and fitness data. In *Proceedings of the 11th EAI international conference on pervasive computing technologies for healthcare*. 163–172. https://doi.org/10.1145/3154862.3154879

[2] Bon Adriel Aseniero, Charles Perin, Wesley Willett, Anthony Tang, and Sheelagh Carpendale. 2020. Activity river: Visualizing planned and logged personal activities for reflection. In *Proceedings of the International Conference on Advanced Visual Interfaces*. 1–9. https://doi.org/10.1145/3399715.3399921

[3] Jillian Aurisano, Abhinav Kumar, Alberto Gonzales, Khairi Reda, Jason Leigh, Barbara Di Eugenio, and Andrew Johnson. 2015. Show me data": Observational study of a conversational interface in visual data exploration. In *IEEE VIS*, Vol. 15. 1.

[4] Tanja Blascheck, Frank Bentley, Eun Kyoung Choe, Tom Horak, and Petra Isenberg. 2021. Characterizing Glanceable Visualizations: From Perception to Behavior Change. In *Mobile Data Visualization*. Chapman and Hall/CRC, 151–176.

[5] Tanja Blascheck, Lonni Besançon, Anastasia Bezerianos, Bongshin Lee, and Petra Isenberg. 2018. Glanceable visualization: Studies of data comparison performance on smartwatches. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 630–640. https://doi.org/10.1109/TVCG.2018.2865142

[6] Niall Bolger, Angelina Davis, and Eshkol Rafaeli. 2003. Diary methods: Capturing life as it is lived. *Annual review of psychology* 54, 1 (2003), 579–616.

[7] Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. 2003. Multimodal'eyes-free'interaction techniques for wearable devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 473–480. https://doi.org/10.1145/642611.642694

[8] Ana Isabel Canhoto and Sabrina Arp. 2017. Exploring the factors that support adoption and sustained use of health and fitness wearables. *Journal of Marketing Management* 33 (1 2017), 32–60. Issue 1-2. https://doi.org/10.1080/0267257X.2016.1234505

[9] Scott Carter and Jennifer Mankoff. 2005. When participants do the capturing: the role of media in diary studies. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 899–908. https://doi.org/10.1145/1054972.1055098

[10] Marta E Cecchinato, Anna L Cox, and Jon Bird. 2017. Always on (line)? User experience of smartwatches and their role within multi-device ecologies. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3557–3568. https://doi.org/10.1145/3025453.3025538

[11] Yang Chen. 2017. Visualizing large time-series data on very small screens. In *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers*. 37–41. https://doi.org/10.2312/eurovisshort.20171130

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

[13] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding self-reflection: how people reflect on personal data through visual data exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 173–182. https://doi.org/10.1145/3154862.3154881

[14] 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. *Conference on Human Factors in Computing Systems - Proceedings*, 1143–1152. https://doi.org/10.1145/2556288.2557372

[15] Stephanie Hui-Wen Chuah. 2019. You inspire me and make my life better: Investigating a multiple sequential mediation model of smartwatch continuance intention. *Telematics and Informatics* 43 (2019), 101245. https://doi.org/10.1016/j.tele.2019.101245

[16] Jaemin Chun, Anind Dey, Kyungtaek Lee, and Seung Jun Kim. 2018. A qualitative study of smartwatch usage and its usability. *Human Factors and Ergonomics In Manufacturing* 28 (7 2018), 186–199. Issue 4. https://doi.org/10.1002/hfm.20733

[17] Karen Church and Barry Smyth. 2009. Understanding the intent behind mobile information needs. In *Proceedings of the 14th international conference on Intelligent user interfaces*. 247–256. https://doi.org/10.1145/1502650.1502686

[18] 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. *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 647–658. https://doi.org/10.1145/2750858.2807554

[19] Sunny Consolvo and Miriam Walker. 2003. Using the experience sampling method to evaluate ubicomp applications. *IEEE pervasive computing* 2, 2 (2003), 24–31. https://doi.org/10.1109/MPRV.2003.1203750

[20] Coorevits, Lynn and Coenen, Tanguy. 2016. The rise and fall of wearable fitness trackers. In *Academy of Management* (Anaheim). 24.

[21] Kenneth Cox, Rebecca E Grinter, Stacie L Hibino, Lalita Jategaonkar Jagadeesan, and David Mantilla. 2001. A multi-modal natural language interface to an information visualization environment. *International Journal of Speech Technology* 4, 3 (2001), 297–314. https://doi.org/10.1023/A:1011368926479

[22] Mihaly Csikszentmihalyi and Reed Larson. 2014. Validity and reliability of the experience-sampling method. In *Flow and the foundations of positive psychology*. Springer, 35–54. https://doi.org/10.1007/978-94-017-9088-8_3

[23] David Dearman, Melanie Kellar, and Khai N Truong. 2008. An examination of daily information needs and sharing opportunities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 679–688. https://doi.org/10.1145/1460563.1460668

[24] Milad Dehghani. 2018. Exploring the motivational factors on continuous usage intention of smartwatches among actual users. *Behaviour & Information Technology* 37, 2 (2018), 145–158. https://doi.org/10.1080/0144929X.2018.1424246

[25] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems*. 667–676. https://doi.org/10.1145/2598510.2598558

[26] Daniel A Epstein, Tanja Blascheck, Sheelagh Carpendale, Raimund Dachselt, and Jo Vermeulen. 2021. Challenges in Everyday Use of Mobile Visualizations. In *Mobile Data Visualization*. Chapman and Hall/CRC, 209–240.

[27] Clayton Feustel, Shyamak Aggarwal, Bongshin Lee, and Lauren Wilcox. 2018. People like me: Designing for reflection on aggregate cohort data in personal informatics systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–21. https://doi.org/10.1145/3264917

[28] Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. 2018. Activity tracking in vivo. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April. https://doi.org/10.1145/3173574.3173936

[29] Rúben Gouveia, Fábio Pereira, Evangelos Karapanos, Sean A Munson, and Marc Hassenzahl. 2016. Exploring the design space of glanceable feedback for physical activity trackers. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*. 144–155. https://doi.org/10.1145/2971648.2971754

[30] 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

[31] Jonna Häkkilä, Farnaz Vahabpour, Ashley Colley, Jani Väyrynen, and Timo Koskela. 2015. Design probes study on user perceptions of a smart glasses concept. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. 223–233. https://doi.org/10.1145/2836041.2836064

[32] Gabriella M Harari, Nicholas D Lane, Rui Wang, Benjamin S Crosier, Andrew T Campbell, and Samuel D Gosling. 2016. Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science* 11, 6 (2016), 838–854. https://doi.org/10.1177/1745691616650285

[33] Javier Hernandez, Daniel McDuff, Christian Infante, Pattie Maes, Karen Quigley, and Rosalind Picard. 2016. Wearable ESM: differences in the experience sampling method across wearable devices. In *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services*. 195–205. https://doi.org/10.1145/2935334.2935340

[34] Stephen Intille, Caitlin Haynes, Dharam Maniar, Aditya Ponnada, and Justin Manjourides. 2016. μEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 1124–1128. https://doi.org/10.1145/2971648.2971717

[35] Alaul Islam, Ranjini Aravind, Tanja Blascheck, Anastasia Bezerianos, and Petra Isenberg. 2022. Preferences and Effectiveness of Sleep Data Visualizations for Smartwatches and Fitness Bands. In *CHI Conference on Human Factors in Computing Systems*. 1–17.

[36] Alaul Islam, Anastasia Bezerianos, Bongshin Lee, Tanja Blascheck, and Petra Isenberg. 2020. Visualizing information on watch faces: A survey with smartwatch users. (9 2020). http://arxiv.org/abs/2009.00750

[37] Alaul Islam, Tanja Blascheck, and Petra Isenberg. 2022. Context Specific Visualizations on Smartwatches. In *EuroVis 2022 - Posters*. Rome, Italy. https://doi.org/10.2312/evp.20221122

[38] Hayeon Jeong, Heepyung Kim, Rihun Kim, Uichin Lee, and Yong Jeong. 2017. Smartwatch Wearing Behavior Analysis. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1 (9 2017), 1–31. Issue 3. https://doi.org/10.1145/3131892

[39] Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A diary study on combining multiple information devices in everyday activities and tasks. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 3903–3912. https://doi.org/10.1145/2702123.2702211

[40] Young-Ho Kim, Diana Chou, Bongshin Lee, Margaret Danilovich, Amanda Lazar, David E Conroy, Hernisa Kacorri, and Eun Kyoung Choe. 2022. MyMove: Facilitating Older Adults to Collect In-Situ Activity Labels on a Smartwatch with Speech. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3491102.3517457

[41] Young-Ho Kim, Bongshin Lee, Arjun Srinivasan, and Eun Kyoung Choe. 2021. Data@Hand: Fostering Visual Exploration of Personal Data On Smartphones Leveraging Speech and Touch Interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 462, 17 pages. https://doi.org/10.1145/3411764.3445421

[42] Ricardo Langner, Lonni Besançon, Christopher Collins, Tim Dwyer, Petra Isenberg, Tobias Isenberg, Bongshin Lee, Charles Perin, and Christian Tominski. 2021. An Introduction to Mobile Data Visualization. In *Mobile Data Visualization*. Chapman and Hall/CRC, 1–32.

[43] Reed Larson and Mihaly Csikszentmihalyi. 2014. The experience sampling method. In *Flow and the foundations of positive psychology*. Springer, 21–34. https://doi.org/10.1007/978-94-017-9088-8_2

[44] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*. https://doi.org/10.1145/1753326.1753409

[45] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding My Data, Myself: Supporting Self-Reflection with Ubicomp Technologies. *Proceedings of the 13th international conference on Ubiquitous computing - UbiComp '11*. https://doi.org/10.1145/2030112.2030166

[46] Xing Liu, Tianyu Chen, Feng Qian, Zhixiu Guo, Felix Xiaozhu Lin, Xiaofeng Wang, and Kai Chen. 2017. Characterizing smartwatch usage in the wild. *MobiSys 2017 - Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, 385–398. https://doi.org/10.1145/3081333.3081351

[47] Jimmy Moore, Pascal Goffin, Jason Wiese, and Miriah Meyer. 2021. An Interview Method for Engaging Personal Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (2021), 1–28. https://doi.org/10.1145/3494964

[48] Suhaib Mujahid, Giancarlo Sierra, Rabe Abdalkareem, Emad Shihab, and Weiyi Shang. 2018. An empirical study of Android Wear user complaints. *Empirical Software Engineering* 23 (12 2018), 3476–3502. Issue 6. https://doi.org/10.1007/s10664-018-9615-8

[49] Ali Neshati, Fouad Alallah, Bradley Rey, Yumiko Sakamoto, Marcos Serrano, and Pourang Irani. 2021. SF-LG: Space-Filling Line Graphs for Visualizing Interrelated Time-series Data on Smartwatches. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*. 1–13. https://doi.org/10.1145/3447526.3472040

[50] Ali Neshati, Bradley Rey, Ahmed Shariff Mohommed Faleel, Sandra Bardot, Celine Latulipe, and Pourang Irani. 2021. BezelGlide: Interacting with Graphs on Smartwatches with Minimal Screen Occlusion. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 501, 13 pages. https://doi.org/10.1145/3411764.3445201

[51] Ali Neshati, Yumiko Sakamoto, and Pourang Irani. 2019. Challenges in displaying health data on small smartwatch screens. *Studies in Health Technology and Informatics* 257 (2019), 325–332. https://doi.org/10.3233/978-1-61499-951-5-325

[52] Ali Neshati, Yumiko Sakamoto, Launa Leboe-Mcgowan, Jason Leboe-McGowan, Marcos Serrano, and Pourang Irani. 2019. G-sparks: Glanceable sparklines on smartwatches. In *45th Conference on Graphics Interface (GI 2019)*. 1–9. https://doi.org/10.20380/GI2019.23

[53] Obi Ogbanufe and Natalie Gerhart. 2018. Watch It! Factors Driving Continued Feature Use of the Smartwatch. *International Journal of Human-Computer Interaction* 34 (11 2018), 999–1014. Issue 11. https://doi.org/10.1080/10447318.2017.1404779

[54] Debajyoti Pal, Anuchart Tassanaviboon, Chonlameth Arpnikanondt, and Borworn Papasratorn. 2020. Quality of Experience of Smart-Wearables: From Fitness-Bands to Smartwatches. *IEEE Consumer Electronics Magazine* 9 (1 2020), 49–53. Issue 1. https://doi.org/10.1109/MCE.2019.2941462

[55] Sabrina Paneels and Jonathan C Roberts. 2009. Review of designs for haptic data visualization. *IEEE Transactions on Haptics* 3, 2 (2009), 119–137. https://doi.org/10.1109/TOH.2009.44

[56] Stefania Pizza, Barry Brown, Donald McMillan, and Airi Lampinen. 2016. Smartwatch in vivo. *Conference on Human Factors in Computing Systems - Proceedings*, 5456–5469. https://doi.org/10.1145/2858036.2858522

[57] Bernd Ploderer, Wolfgang Reitberger, Harri Oinas-Kukkonen, and Julia van Gemert-Pijnen. 2014. Social interaction and reflection for behaviour change. , 1667–1676 pages. https://doi.org/10.1007/s00779-014-0779-y

[58] Aditya Ponnada, Caitlin Haynes, Dharam Maniar, Justin Manjourides, and Stephen Intille. 2017. Microinteraction ecological momentary assessment response rates: Effect of microinteractions or the smartwatch? *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–16. https://doi.org/10.1145/3130957

[59] Reza Rawassizadeh, Chelsea Dobbins, Manouchehr Nourizadeh, Zahra Ghamchili, and Michael Pazzani. 2017. A natural language query interface for searching personal information on smartwatches. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. 679–684. https://doi.org/10.1109/PERCOMW.2017.7917645

[60] John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. 2014. Personal tracking as lived informatics. *Conference on Human Factors in Computing Systems - Proceedings*, 1163–1172. https://doi.org/10.1145/2556288.2557039

[61] Alexander Schiewe, Andrey Krekhov, Frederic Kerber, Florian Daiber, and Jens Krüger. 2020. A Study on Real-Time Visualizations During Sports Activities on Smartwatches. In *19th International Conference on Mobile and Ubiquitous Multimedia* (Essen, Germany) *(MUM '20)*. Association for Computing Machinery, New York, NY, USA, 18–31. https://doi.org/10.1145/3428361.3428409

[62] Grace Shin, Mohammad Hossein Jarrahi, Yu Fei, Amir Karami, Nicci Gafinowitz, Ahjung Byun, and Xiaopeng Lu. 2019. Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review. https://doi.org/10.1016/j.jbi.2019.103153

[63] Carolin Siepmann and Pascal Kowalczuk. 2021. Understanding continued smartwatch usage: the role of emotional as well as health and fitness factors. *Electronic Markets* (2021). https://doi.org/10.1007/s12525-021-00458-3

[64] Timothy Sohn, Kevin A Li, William G Griswold, and James D Hollan. 2008. A diary study of mobile information needs. In *Proceedings of the sigchi conference on human factors in computing systems*. 433–442. https://doi.org/10.1145/1357054.1357125

[65] Arjun Srinivasan, Bongshin Lee, Nathalie Henry Riche, Steven M. Drucker, and Ken Hinckley. 2020. InChorus: Designing Consistent Multimodal Interactions for Data Visualization on Tablet Devices. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3313831.3376782

[66] Arjun Srinivasan, Nikhila Nyapathy, Bongshin Lee, Steven M Drucker, and John Stasko. 2021. Collecting and characterizing natural language utterances for specifying data visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–10. https://doi.org/10.1145/3411764.3445400

[67] Arjun Srinivasan and John Stasko. 2017. Natural language interfaces for data analysis with visualization: Considering what has and could be asked. In *Proceedings of the Eurographics/IEEE VGTC conference on visualization: Short papers*. 55–59. https://doi.org/10.2312/eurovisshort.20171133

[68] Niels Van Berkel, Denzil Ferreira, and Vassilis Kostakos. 2017. The experience sampling method on mobile devices. *ACM Computing Surveys (CSUR)* 50, 6 (2017), 1–40. https://doi.org/10.1145/3123988

[69] Aku Visuri, Zhanna Sarsenbayeva, Niels Van Berkel, Jorge Goncalves, Reza Rawassizadeh, Vassilis Kostakos, and Denzil Ferreira. 2017. Quantifying sources and types of smartwatch usage sessions. *Conference on Human Factors in Computing Systems - Proceedings* 2017-May, 3569–3581. https://doi.org/10.1145/3025453.3025817

[70] Aku Visuri, Niels van Berkel, Jorge Goncalves, Reza Rawassizadeh, Denzil Ferreira, and Vassilis Kostakos. 2021. Understanding usage style transformation during long-term smartwatch use. *Personal and Ubiquitous Computing* (2021). https://doi.org/10.1007/s00779-020-01511-2

[71] Ryan Vooris, Matthew Blaszka, and Susan Purrington. 2019. Understanding the wearable fitness tracker revolution. *International Journal of the Sociology of Leisure* 2 (12 2019), 421–437. Issue 4. https://doi.org/10.1007/s41978-018-00022-y

[72] Xinghui Yan, Shriti Raj, Bingjian Huang, Sun Young Park, and Mark W Newman. 2020. Toward Lightweight In-situ Self-reporting: An Exploratory Study of Alternative Smartwatch Interface Designs in Context. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–22. https://doi.org/10.1145/3432212

Received 15 May 2022; revised 15 August 2022; accepted 10 October 2022