Understanding Quantified-Selfers' Practices in Collecting and Exploring Personal Data

Eun Kyoung Choe[1], Nicole B. Lee[2], Bongshin Lee[2], Wanda Pratt[1], Julie A. Kientz[1]

[1]University of Washington
{eunky, wpratt, jkientz}@uw.edu

[2]Microsoft Corporation
{nilee, bongshin}@microsoft.com

ABSTRACT

Researchers have studied how people use self-tracking technologies and discovered a long list of barriers including lack of time and motivation as well as difficulty in data integration and interpretation. Despite the barriers, an increasing number of *Quantified-Selfers* diligently track many kinds of data about themselves, and some of them share their best practices and mistakes through Meetup talks, blogging, and conferences. In this work, we aim to gain insights from these "extreme users," who have used existing technologies and built their own workarounds to overcome different barriers. We conducted a qualitative and quantitative analysis of 52 video recordings of Quantified Self Meetup talks to understand what they did, how they did it, and what they learned. We highlight several common pitfalls to self-tracking, including tracking too many things, not tracking triggers and context, and insufficient scientific rigor. We identify future research efforts that could help make progress toward addressing these pitfalls. We also discuss how our findings can have broad implications in designing and developing self-tracking technologies.

Author Keywords

Quantified Self; self-monitoring; self-tracking; health; personal informatics; personal analytics; self-experimentation.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User-centered design; J.3. Life and medical sciences: Health.

INTRODUCTION

Although many people do not routinely track personal data, *Quantified-Selfers* (Q-Selfers) are notable exceptions who diligently track many kinds of data about themselves. They are a diverse group of life hackers, data analysts, computer scientists, early adopters, health enthusiasts, productivity gurus, and patients. Believing in the notion of "self-knowledge through numbers," Wired Magazine editors Gary Wolf and Kevin Kelly created a blog called quantifiedself.com in 2007, which has become the repository for people to share self-tracking practices. At the core of Quantified Self (QS) are the frequent, in-person grassroots Meetups (meetings) where participants share their best practices, experiences, and mistakes.

On the academic side, human-computer interaction researchers and designers have developed and studied many self-tracking technologies in the domain of health and wellness [13,27]. Similar to the QS movement, the field of Personal Informatics (or Personal Analytics) adopts the approach that through knowledge of one's data, it becomes possible to reflect on one's activities, make self-discoveries, and use that knowledge to make changes. Although researchers acknowledged the value of self-tracking technologies (e.g., [2,4,19,20,23]), they also discovered a long list of barriers toward the adoption of self-tracking technologies. These barriers included lack of time, insufficient motivation, unsuitable visualization and analytics tools, poor skills for analyzing data, and fragmented data scattered across multiple platforms [17].

In light of the barriers they face, Q-Selfers offer us a useful perspective from which to re-examine the current design of self-tracking technologies and ways to improve them. Q-Selfers encompass a broad spectrum of people ranging from those who use pen and paper to those who build their own tracking applications. Because they can be categorized as a somewhat *extreme user group*, their stories, including the successful ones, might not be generalizable or applicable to the broader population. However, as other researchers point out [29], the perspective of those who represent "extremes" gives us distinct insights because they have used existing technologies and spent numerous hours building their own workarounds when faced with problems.

We had many questions about this particular group. What motivates Q-Selfers to keep tracking data, despite numerous barriers? What tools do they use to collect and explore data? What insights do they gain from tracking? What are the outcomes of tracking? What challenges do they face and how do they overcome these? We explored these questions through a qualitative and quantitative analysis of 52 video recordings of QS Meetup talks stored on the QS blog [25]. Each talk illustrates a distinctive self-tracking approach that

CHI 2014, April 26 – May 01 2014, Toronto, ON, Canada

ACM 978-1-4503-2473-1/14/04$15.00.
http://dx.doi.org/10.1145/2556288.2557372

could benefit human-computer interaction, health informatics, and information visualization researchers whose work is within the domain of self-tracking and personal analytics.

In what follows, we present the results from our study on Q-Selfers' practices of collecting and exploring their personal data. We begin by providing background on self-monitoring and the rise of Quantified Self movement. Next, we explain our study methods, dataset, and profiles of Q-Selfers. We then detail themes that arise from our qualitative and quantitative analysis as we answer the Three Prime Questions posed to the QS Meetup speakers—(1) what they did, (2) how they did it, and (3) what they learned. While addressing these questions, we highlight several common pitfalls Q-Selfers experience. These pitfalls include tracking too many things, not tracking triggers and context, and lacking scientific rigor. Q-Selfers offer workarounds in addressing some of these issues, but the questions of how to easily explore data and how to bring scientific rigor to the Quantified-Self movement remain open and require further research. We identify future research efforts that could help make progress toward addressing these issues.

RELATED WORK

In this section, we provide some background on self-monitoring, the rise of Quantified Self movement, and scholarly and business endeavors in this space.

Self-Monitoring (or Self-Tracking)

Although using technology to monitor one's own behavior is a relatively new concept, self-monitoring (or self-tracking)—the process of recording one's own behaviors, thoughts, and feelings—is an area of research within behavioral psychology, which dates back to 1970 [14]. Self-monitoring has been traditionally employed in clinical and research settings to serve an *assessment* function as well as a part of *treatment* function within behavior therapy [15]. The role of clinicians was important in self-monitoring—they used self-monitoring for all stages of assessment, such as diagnosis, target behavior selection for treatment, functional assessment, and treatment monitoring.

More recently, self-monitoring has been widely embodied in the design of sensing and monitoring applications because of its effectiveness on increased awareness and behavior change. Sensors have become smaller and better integrated with mobile devices, making it easy for people to track numerous types of data. Recognizing the power of self-monitoring in promoting health behavior change, researchers and designers often incorporate automated sensing or manual tracking feature in designing self-monitoring technology. Within the health domain, researchers and companies designed technology for tracking physical fitness (e.g., [4,6,9,16,19,22]), sleep (e.g., [6,11,16,31]), diet [20], smoking [1], and stress [21]. In addition, the Mobile Health Mashups system shows significant correlations across sensor data from multiple sources such as exercise, weight, food, sleep, and mood [2].

Tracking a health indicator or symptom has become popular among the general public. In a nation-wide survey on peoples' health tracking practice, researchers showed that seven in ten U.S. adults track a health indicator for themselves or for a loved one [7]. However, among those who track one or more health indicators, only 21% use some form of technology for tracking while 49% keep track of progress "in their head" and 34% track data on paper [7]. In the same survey, researchers found that people with chronic conditions are significantly more likely to track a health indicator or symptom. Mobile phones have become a promising platform for patients to do health tracking because of their ability to support journaling, text messaging, and automated sensing [13]. Through qualitative inquiry of cancer patients' symptom tracking technology use, Patel and colleagues revealed how patient-led information capture and management could help patients feel psychosocial comfort, be prepared for the attending rounds, and improve symptom communication with clinicians [23].

Quantified Self and Personal Analytics

Quantified Self (QS) refers to the name of a community as well as the practice of self-tracking. The prevalence of low cost monitoring sensors accelerated the rise of the Quantified Self movement [25]. Initially started in the Silicon Valley area among technology enthusiasts, QS has become a community of people practicing self-monitoring and building self-monitoring technology. QS as a community promotes sharing of individual self-tracking practices through Meetups, blogging, and annual conferences. As of January 2014, QS is an active, international community, with Meetups held in 106 cities in 36 countries. They have held an annual conference since 2011. Identifying health tracking as a promising area for growth, toolmakers of self-monitoring devices and software attend Meetups to promote their products and sponsor the annual QS conference.

QS goes by other terms, such as personal analytics [28] and personal informatics [17]. These all refer to a class of systems or practices that help people collect and reflect on personal information. Stephen Wolfram, a creator of software Mathematica and of Wolfram Alpha (a knowledge engine) as well as an avid self-tracker himself, coined the term personal analytics and now applies analytical techniques to people's personal data [30]. His company recently deployed Personal Analytics for Facebook, which automatically analyzes and generates a report on personal relationship and other behaviors on the site.

Li and colleagues coined the term personal informatics, proposed a stage-based model of personal informatics systems composed of five stages (preparation, collection, integration, reflection, and action), and identified barriers people have in each of the stages [17]. In a follow-up paper, they explored how ubiquitous computing technologies could properly support the *self-reflection* stage [18]. Although we share similar goals to understand self-trackers' practice, our work differs in three regards. First, we attempt to learn self-tracking practices from the extreme user group—Quantified-Selfers who volunteered to give a talk in front of other Q-Selfers and share success and even failure stories. Second, we detail how Q-Selfers *explore* data, which encompasses both data analysis and visualization. To better understand how Q-Selfers explored data, we analyzed how they gained insights and what visualizations they created. Lastly, we created profiles of Q-Selfers to characterize their backgrounds and motivations for tracking. In conducting our study, we strived to learn Q-Selfers' common pitfalls and their workarounds to avoid those pitfalls.

THREE PRIME QUESTIONS AND STUDY METHOD

At the QS Meetups, people talk about their firsthand experiences with self-tracking methods and tools using a "Show & Tell" format. Talks on scientific theories, demos of tools and apps, and philosophical speculation are discouraged unless they are grounded in actual attempts at self-tracking and self-experimentation. The uniqueness of the QS Show & Tell comes from the fact that they follow the specific guideline, provided beforehand, to organize the talk to answer the Three Prime Questions:

- What did you do?
- How did you do it?
- What did you learn?

The consistent structure of the talks makes them a valuable dataset. In answering *what they did*, speakers talk about initial problems and motivations to do self-tracking and track items. In answering *how they did it*, they talk about tools and methods they used and the visualizations they created from personal data. In answering *what they learned*, they talk about insights gained and the outcomes of their tracking. The talks are usually 5 to10-minute-long followed by a question and answer period. The talks, including the Q&A, are often video-recorded and uploaded to the quantifiedself.com blog by Meetup organizers for sharing.

Dataset

As of April 8, 2013, 205 video posts had been uploaded to the QS blog since 2008. Of those, we analyzed 83 recent video posts, those uploaded since January 2012, to examine the most up-to-date landscape of QS practice (Figure 1).

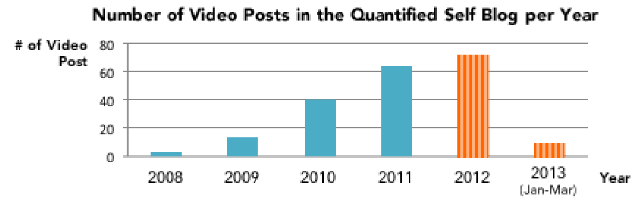

Figure 1. QS Video posts per year. Our dataset is colored in orange with vertical stripes.

Not all of the posts fit within our area of interest for this study. For example, despite the three prime questions guideline, some speakers presented a new tool that they developed without describing their actual self-tracking practice. Others presented academic research on other people's data. Some videos failed to capture the speakers' visual aids (i.e., slides) so it was difficult to understand the specific context. Thus, video posts had to meet the following two inclusion criteria in order to be added to our dataset: (1) the speaker should present their own QS practices; and (2) video posts should include personal data visualizations of some kind (e.g., table, graph) created with the speaker's own data. Of the 83 videos we reviewed, 52 videos met the inclusion criteria. The average length of these videos was 15 minutes, 53 seconds (including Q&A). We transcribed this entire corpus of videos to aid with analysis.

The speakers of QS Meetup talks were a self-selected group of people who volunteered to give a talk, and might not accurately represent the whole community of Q-Selfers, not to mention the general public. We suspect that they are more extreme in terms of technical ability and experience with self-tracking than typical Q-Selfers. We also note that not all the QS Show & Tell videos are recorded and uploaded to the QS blog.

Analysis Techniques

We employed a variety of techniques to analyze our dataset. First, with an aim to understand and characterize Q-Selfers, we created a profile for each speaker by systematically capturing the following information: location, gender, job description, health condition, types of data collected, self-tracking duration, data collection tool, data exploration tool, type of tool (user-generated vs. commercial), and data sharing aspects. Given that we had to rely on the information speakers disclosed during the talk, some information is missing for some speakers. Second, we conducted an affinity analysis [3] as a group (Figure 2). After several passes, the video transcripts were broken into approximately 400 quotes, each of which contained one main idea. We inductively organized these into categories to identify key themes. We drew several bottom-up themes regarding people's motivations for self-tracking, self-tracking methods and tools, insights gained, outcomes of tracking, and common pitfalls of self-tracking. Lastly, we captured 188 screenshots (and the slides when available) that included personal data visualizations. We analyzed the screenshots by categorizing the visualization type.

Figure 2. Affinity analysis of the video transcript quotes.

PROFILES OF THE QUANTIFIED-SELFERS

Here we provide the profiles of Q-Selfers based on qualitative and quantitative coding analysis.

Location. Seventeen (33%) video posts were recorded from San Francisco / Mountain View / Silicon Valley area. This location is where the QS movement first started and still remains very active. Nine (17%) video posts were from Seattle followed by seven (13%) videos from London and New York respectively. Other locations included Toronto (6%), Pittsburgh (4%), Singapore (4%), Washington DC (4%), Boston (2%), San Diego (2%), and Portland (2%).

Gender. Forty-one (79%) speakers were male, while only 11 (21%) were female. Pew Research reports that in the general population, men and women are equally likely to report tracking their weight, diet, or exercise routine [7].

Health Condition. Eighteen (35%) reported having some health conditions, such as sleep disorders, diabetes, panic attacks, cancer, obesity, or allergies. For them, their health conditions highly influenced what they tracked because they wanted to maintain a certain condition, find triggers, identify a medication's effect, or achieve some health goal.

Job Description. Twenty-one (40%) speakers were working at a startup. Eighteen (37%) speakers described themselves as a software engineer or programmer. Seven (13%) were working in data analytics and four (8%) were electrical engineers. Other job titles included creative director, psychologist, designer, product manager, graduate student, operations analyst, professor, and professional athlete.

Tracking Duration. The average duration of tracking was 25 months, where the obtained range was 5 days–20 years (*SD* = 44.0 months, *Median* = 8 months).

WHAT DID YOU DO?

We organize our findings based on the Three Prime Questions. We begin by answering the first question—"what did you do?" We here describe the types of data Q-Selfers tracked and the motivations behind tracking. We also identify common pitfalls regarding data collection phase and Q-Selfers' approaches to alleviate some of the pitfalls.

Figure 3. Number of people tracking a certain item.

Items Tracked

Activity (40% of Q-Selfers), food (31%), weight (29%), sleep (25%), and mood (13%) were the most popular items Q-Selfers reported tracking. In contrast to our results, bank statements, email history, and credit card bills were the top 3 items people reported in Li et al. [17]. This is possibly because Q-Selfers did not consider readily available data "self-tracking," or if they did, they did not report on it during the Meetup talks. On average, Q-Selfers track 2.92 items (*SD* = 2.41) where the obtained range was 1–11. In all, they reported 57 unique items. The long-tail shape of Figure 3 indicates that Q-Selfers have diverse interests. Other items they reported tracking include cognitive performance, blood glucose, location, heart rate, symptoms, knowledge, stress, body fat, productivity, snoring, movies, posture, medicine, skin condition, home energy usage, clothes, and public transit usage. Some people track multiple items simultaneously with the intention of identifying correlations among the factors, while others track one or two items at a time but apply the self-tracking practice on several topics over time as their interests change.

Motivations to Practice Self-Tracking

We classified Q-Selfers' motivations to track into three main categories: (1) to improve health, (2) to improve other aspects of life, and (3) to find new life experiences. In Table 1, we break down these categories further and include tracking examples for each of the categories. Thirty-five (67%) speakers tracked one or more health-related items with an aim to improve aspects of health. Considering the number of people who had a health condition (35%), improving health was a prevalent motivation regardless of the presence of a health condition. Furthermore, they had very

Motivations	Sub-categories	Tracking example
To improve health	To cure or manage a condition	Track blood glucose to hit the target range [P37]
	To achieve a goal	Track weight to get back to the ideal weight of 135 pounds [P39]
	To find triggers	Log triggers that cause atrial fibrillation [P55]
	To answer a specific question	Track niacin intake dosage and sleep to identify how much niacin to take for treating symptoms [P76]
	To identify relationships	Track exercise, weight, muscle mass, and body fat to see the relationships among the factors [P31]
	To execute a treatment plan	Log food, exercise, and panic as a recovery plan for panic attack [P35]
	To make better health decisions	Record ideas of things that thought were healthy and unhealthy to make better decisions [P18]
	To find balance	Log sleep, exercise, and time to get back from erratic lifestyle [P23, P42, P54]
To improve other aspects of life	To maximize work performance	Track time to know the current use of time and ways to be more efficient [P43, P63]
	To be mindful	Take a self-portrait shot everyday for 365 days to capture each day's state of mind [P26]
To find new life experiences	To satisfy curiosity and have fun	Log the frequency of "puns" to see how often these puns happened and what triggered them [P12]
	To explore new things	Track every street walked in Manhattan to explore as much of the city as possible [P34]
	To learn something interesting	Track heart rate for as long as possible and see what can be learned from it [P62]

Table 1. Quantified-Selfers' tracking motivations and examples for each category.

specific health-related goals—such as finding triggers for an allergy, finding out how exercise affects body mass and weight, finding the right drug dosage, or executing a treatment plan for treating panic attacks—rather than merely "to become healthy" or "to change health behaviors." Some people in this group claimed that they "treated" or "cured" a disease through self-tracking. For those who experienced positive outcomes from self-tracking, QS was an *approach* to better life, not just a data collection method.

Another group of Q-Selfers was interested in improving other aspects of life—predominantly work efficiency and cognitive performance. They used self-tracking to measure their current use of time (with time tracking apps or calendar logging), cognitive performance (by taking an online cognitive test), or time spent on a computer (with productivity tracking software). People in this group—who were either software engineers or students—wanted to find ways to "optimize" their work and life and "maximize" learning.

The last category consists of those who wanted to have new life experiences through self-tracking. They often had no specific goals in mind when starting to track, but quickly discovered interesting patterns from data that led to data collection becoming habitual. For example, P62, who did not have a heart condition, collected heart rate data for 24 hours a day for over a year. He streamed his heart rate data to various channels online, which were updated every 30 minutes. He learned how his body responds to various routines and stressful events, which in return influenced his decision-making (e.g., avoiding heavy meals because his heart rate would go up 20%). The technical ability of people in this group combined with their creativity allowed them to explore new life experiences through self-tracking.

Common Pitfall 1: Tracking Too Many Things

Q-Selfers described that they were often too ambitious at first and tried to track too many things, as P61 remarked: "*I can honestly say that I've made the classic newbie self-tracking mistake which is that I track everything*." Tracking too many things often led to either stop tracking entirely due to *tracking fatigue* or failure to do data analysis due to too much data in different formats.

Q-Selfers offered some suggestions on how to alleviate tracking fatigue. First, they suggested automating the tracking and data uploading if possible. P39 had been tracking her weight, food, and exercise for 6 years using several different tracking methods, such as manual entry with an Excel spreadsheet, pen and paper, Google docs, and most recently, automatically through a WiFi-scale. Her definition of successful tracking was to capture many data points for a long period of time. The use of a WiFi-scale, which automatically uploads data to a website, allowed her the longest and the most regular data acquisition (compared to other methods). Second, if automating is not possible, Q-Selfers suggested making tracking very simple and easy to do by (1) lowering data granularity (e.g., "*If you can't automate your tracking, make your tracking binary*" [P51]) or (2) making manual capture very easy. P35 built a manual counter app whose main design goal was to reduce user burden in capturing his panic symptoms and triggers: "*...recording is a one-tap process—have a drink at the bar, tap alcohol, go on a run, tap to start, tap to stop, simple*." At the expense of data granularity, they were able to lower the user burden associated with capturing, thereby capturing more data points overall. Lastly, Q-Selfers suggested making tracking a *rewarding experience*. P11 drew an interesting analogy to explain what tracking means to her: "*...when I was pretty young, I was really susceptible to being awarded Gold Stars, it makes me want to do the thing I've been awarded Gold Star for more. So the process of tracking was like awarding myself the Gold Star... So what I learned was track what you want to do more.*" P11 explained that focusing on the positives makes the tracking experience rewarding and less of a burden.

Common Pitfall 2: Not Tracking Triggers and Context

People who were new to QS made a common mistake of focusing too much on tracking symptoms or outcome measures but failing to capture the important triggers or context. This failure resulted in not having enough clues on how to improve outcome measures. P9 described, "*...I've been trying all this biometric tracking trying to be more consistent in my health than have more healthy habits. But the whole time, not just my health habits, but even my tracking habits were completely reliant on my emotional state. So here I was trying to track all these symptoms, and I was completely ignoring the cause*." After a few months of trial and error, P9 modified her tracking routine from capturing biomedical data to capturing negative emotion and the biometric data surrounding it. Likewise, P3—a student who diligently tracked every activity—was able to alter his initial question after a few months of tracking. He initially wanted to know *where his time was going* as specifically as possible. He kept track of time and planned everything ahead using a calendar. However, after tracking for four months, he realized he needed to step back from the day-to-day events and ask a different question: "*how to balance my life?*" It is difficult to know exactly what to track or what questions to ask in the beginning. In fact, the initial tracking phase helped Q-Selfers redefine what to track or what questions to ask. Q-Selfers thus endeavored to step back from time-to-time and reflect on whether they are tracking the right thing for the right reason. This finding is in line with Li et al.'s report on the transitions between the Maintenance phase and the Discovery phase during self-reflection [18].

HOW DID YOU DO IT?

In this section, we address the second prime question, "*how did you do it?*" We examine tools for data collection and exploration, reasons for building custom tools, and visualizations Q-Selfers created. We also describe the notion of *self-experimentation*, a prevalent practice among Q-Selfers to get concrete answers to their questions.

Data Collection and Exploration Tools

Q-Selfers reported using a variety of tools for self-tracking, which we categorize into *Data Collection Tools* and *Data Exploration Tools* (see Table 2). Data exploration tools include data analysis and visualization tools. On average, Q-Selfers used 2.1 data collection tools (SD = 1.08) and 1.4 data exploration tools (SD = 0.63).

Data Collection Tools. Commercial hardware, such as a health monitoring device (e.g., Fitbit, ZEO, WIFI-scale, heart rate monitor), was the most popular tool (56%) followed by spreadsheets, such as Excel or Google Docs (40%). Eleven (21%) built custom software such as a snoring app, mood/stress tracking app, activity/location tracking app, or productivity tracking software. Ten (19%) used commercial software, such as standalone mobile apps for tracking sleep, productivity, or food. Two speakers reported building custom hardware, such as wearable sensors for tracking posture and smiles.

Data Exploration Tools. The most popular data exploration tool was a spreadsheet (e.g., Excel, Google spreadsheet) for running simple statistics and creating graphs (44%). Eighteen (35%) built custom software that required some programming such as using open-source JavaScript libraries to create a website or mobile apps with data visualization features. Fourteen (27%) relied on a commercial website (e.g., Fitbit, ZEO, Quantified Mind) where visualizations are auto-generated once data is manually entered or uploaded from commercial hardware. Six (12%) used commercial software that often interconnected with commercial hardware to aid with analytics and data storage (e.g., software that comes with blood glucose monitor). Only two speakers used statistical software, such as R. None mentioned using commercial data exploration software (e.g., Tableau).

In all, thirty (58%) used only commercial tools, twelve (23%) used only user-generated (custom) tools, and ten (19%) used a mix of commercial and user-generated tools for collecting and exploring data. Our analysis shows that not many tools support the whole spectrum of QS from data collection to data exploration. We also found that many Q-Selfers built custom tools, especially for data exploration purposes. Both findings indicate issues with data portability, regarding people having to export and import data from a tracking tool to an exploration tool. To further exacerbate the situation, some companies (e.g., Fitbit) charge fees for people to export their data, which makes it hard to combine data from different sources.

Data Collection Tool	% (#)
commercial hardware	56% (29)
spreadsheet	40% (21)
custom software	21% (11)
pen and paper	21% (11)
commercial software	19% (10)
commercial website	10% (5)
camera	6% (3)
open-source platform	6% (3)
custom hardware	4% (2)
other	10% (5)

Data Exploration Tool	% (#)
spreadsheet	44% (23)
custom software	35% (18)
commercial website	27% (14)
commercial software	12% (6)
open-source platform	8% (4)
statistical software	4% (2)
pen and paper	2% (1)

Table 2. Types of data collection tools and data exploration tools and usage frequency.

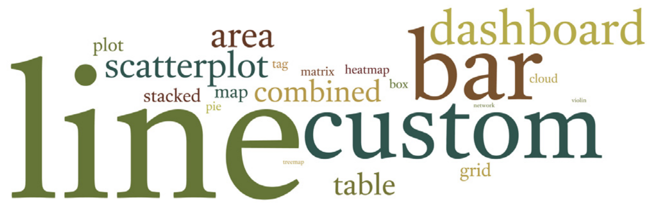

Figure 4. Tag cloud showing the usage frequency of visualization types. Line chart, bar chart, and custom visualizations were the top 3 most commonly used ones.

Visualization Types

Visualizations were the key means to gain insights from data. To analyze how Q-Selfers explored data, we captured 188 screenshots composed of 243 charts and types of graphical feedback. From these, we analyzed visualization types and frequency of usage. We identified 21 unique visualization types, which are shown in Figure 4 (word size reflects the frequency of usage).

Line charts were by far the most frequently used, followed by bar charts and custom visualizations such as an infographic-style website, calendar (Figure 5-a), physical light (Figure 5-b), map and photo grid on timeline (Figure 5-c), and a combined visualization composed of line charts, stacked bar charts, and tables (Figure 5-d). Figure 5-a is an example of "*appropriation*" where iCal, a personal calendar application, was used for tracking and visualization purposes. Figure 5-b was a rare example of real-time feedback—the blue lights blink whenever wearable EEG sensors detect smiling. Figure 5-c and Figure 5-d were highly customized and complex timeline visualizations, which helped creators understand how they spent time online and offline (Figure 5-c) and how to optimize performance (Figure 5-d).

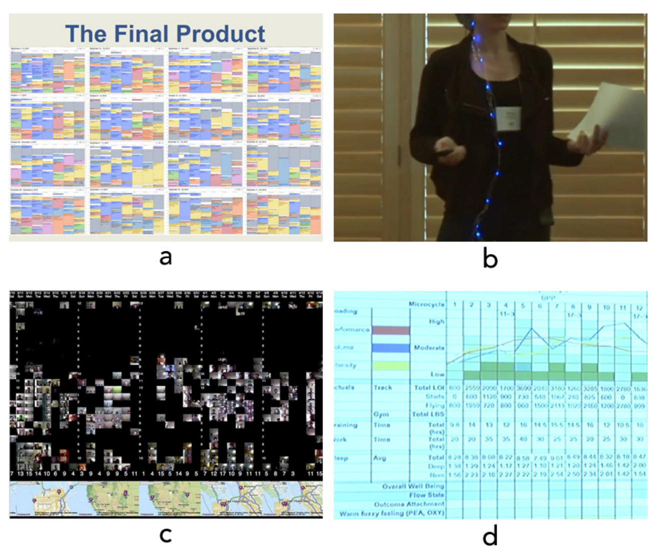

Figure 5. Examples of custom visualizations.

Reasons for Building Custom Tools

Although numerous commercial self-tracking tools are available, many Q-Selfers built their own tools. We identified common reasons for building custom tools. First, few commercial tools support the two key features that Q-Selfers prefer—(1) being able to track and explore data using a *single tool*, and (2) being able to perform *self-experimentation*. When Q-Selfers had the technical ability, they built a custom tool to meet their needs. For example, P15 had a snoring problem. He first looked for existing snoring apps, but they did not have the features he wanted. He envisioned an app that could do both snore tracking and analysis. He also wanted to test if certain things (e.g., snoring remedies, drug, alcohol) affected his snoring. Not finding what he wanted, he built an app called SnoreLab, which he released to a commercial app store. Second, Q-Selfers built a new tool when they wanted to do centralized tracking as P23 remarked, "*I found myself using Nike Plus for my exercise, my Foursquare for social check-ins and several different apps just for tracking my time. When I finally said, 'You know what? I'm fed up with this. I want to make my own tool that allows me to do this in a more cohesive manner.'*" Third, many Q-Selfers built custom websites for data presentation, which was typically done with publicly available visualization APIs such as d3 [5] and the Google Charts API [8]. Lastly, some Q-Selfers built a custom tool simply because no existing tool supported their needs. One example is software developed by P70 who tracked her inventory of clothes that would help her coordinating clothes and simplify her wardrobe.

Self-Experimentation

Q-Selfers wanted to draw definitive conclusions from their QS practice—such as identifying correlation (e.g., sleep and cognitive performance are not correlated) or even causation (e.g., weight tracking causes weight loss). To accomplish this goal, they needed to first generate hypotheses to test. Testing ideas came from careful observations of previous behavioral patterns (e.g., P39 strongly suspected a beer allergy), other speakers' Show & Tell talks (e.g., P33 was inspired by another Q-Selfer who tested whether eating butter increases cognitive performance), or individual needs (e.g., P76 wanted to find the right medication dosage).

Q-Selfers often described the process of seeking answers as *self-experimentation*. When used in an academic context, self-experimentation means participating in one's own experiments when recruiting other participants is not feasible. However, in QS, the goal of self-experimentation is *not* to find generalizable knowledge, but to find meaningful self-knowledge that matters to individuals. P33 emphasized, "*I discovered the importance of testing for myself because what works for some people does not work for me*."

Common Pitfall 3: Lack of Scientific Rigor

Q-Selfers conducted a wide variety of self-experimentations without having a control condition. A typical personal experiment resembles the experiment that P69 conducted—he observed patterns between his allergic reactions and spending on beer, so he suspected that beer might have caused his allergy: "*So ignoring the doctor, who said I was fine, I decided to do one final experiment. We had colleagues going away, and it was a very happy day, I drank lots of beers. By the following Sunday and the following Monday I was in the worst form that I had ever been, and I decided that was enough for me. This must be the trigger.*" P69 did not control for other confounding factors, which threatens the internal validity of his finding. However, as long as Q-Selfers were happy with the outcomes of tracking, most of them did not seem to care about the lack of scientific rigor, as it is not the main goal of QS.

Although a minority, some Q-Selfers attempted to design more rigorous personal experiments. We identified three approaches that could possibly increase internal validity—(1) having a control condition, (2) triangulating with other methods, and (3) using the experience sampling method. We note that the terms we used in this paper such as "internal validity" and "triangulation" are our interpretation of the behavior, not the language used by the Q-Selfers.

Some Q-Selfers conducted self-experimentation with a control condition to reduce biases. As a doctor and researcher, P55 was well aware of skepticism his colleagues had about QS. He said, "*When I talk about these things, I feel like I'm talking by myself because most of all in medicine, and in science, they're going to roll their eyes. It's like science is being done in your garage, but I really think that there's some real potential*." Then he explained how he conducted a within-subjects design to identify what triggers his atrial fibrillation by comparing what he did right before the onset of the disease (hazard period) to the usual routine (control period). Then he calculated an Odds Ratio—a measure of association between an exposure and an outcome—and identified risk factors for his atrial fibrillation such as caffeine, air flight stress, more than 1 glass of wine, and public speaking in the previous 2 hours of the onset of the disease.

Triangulation—using two or more different methods to measure the same phenomenon—was commonly used to facilitate data validation through cross verification from multiple sources. P28 said, "*... if I compare my Zeo data to my Fitbit data I really only wake up when I flip over in bed, so it's actually very accurate for me.*" However, P7 came up with a disturbing finding: "*I discovered that my glucose meters aren't that good. So, comparing the measurements from these two different meters, I only came out with R-squared of 0.46, which I would have hoped for a lot better agreement between the two meters.*"

Some Q-Selfers employed the *experience sampling method* (ESM). P59 used an app called 'The Mappiness' to track his stress level. He configured the app to prompt him at random times during the day. The number of prompts was also configurable. He acknowledged that ESM produces the gold standard of experience measurement.

Nevertheless, critics abound. Although some biases might be reduced from some of these attempts, critics claim that *experimenters* might be biased to produce the result they expect to see [26]. By definition, QS is designed and conducted by the experimenter, and thus, the issue regarding experimenter's bias remains open.

WHAT DID YOU LEARN?

The way Q-Selfers reported their learning was twofold. First, they reported insights gained from their data exploration. Second, they reported desirable and undesirable outcomes in a broader tracking context. After we report on these two types of learning, we discuss the key hurdle in gaining insights, which is data interpretation.

Gained Insights

Q-Selfers often summarized their findings by reporting *descriptive statistics*. For example, after tracking the usage of tabs in a web browser, P4 learned that in a two-month period he opened or closed 32,000 tabs, which averages to about 500 a day. This finding led him to think about his next project: extending his code to track every time he switches tabs as a proxy for how much attention he is giving to something. *Comparison measures* were common descriptive statistical methods that helped Q-Selfers quickly gain insights. When both control and intervention conditions were in place, they typically reported differences in means or Odds ratio. However, they could still compare within themselves without the explicit control condition by categorizing data points in several bins after data collection was complete and then comparing differences across the bins. One example is how P78 reported his sleep data—he compared sleep efficiency and time to fall asleep across restless nights and restful nights. Q-Selfers also compared themselves against the general population with similar demographics if they had access to the population data. P70, a Canadian engineer who tracked work time, learned that she worked more hours than average for Canadian workers.

Q-Selfers with a statistical background reported statistical test results. *Correlation* was the most commonly reported statistical test. For example, P18 learned that "*high idea days were correlated with conferences, sedentary events, Internet usages, high calories, and very little working out*," which he found problematic. People were surprised when they found low correlation between things that they thought were highly associated: "*So far, I had pretty much no correlation, so it's really interesting to think about how I have found no correlations between even the most meaningful things in my life and how I rate the day*" [P17]. Q-Selfers with no statistical background still reported perceived correlation in layperson's language: "*When my symptoms were good or when my body felt good, I happened to be in a good mood. When my symptoms were bad or my body felt bad, the mood or my mental state was bad*" [P61].

Stepping back from the data, Q-Selfers articulated high-level, qualitative take-away points. For example, P11 learned that "*the biggest contributors to my daily happiness are the small things*," and P26 learned "*how tragic it is that we all age*" after taking self-portrait shots every day for a year. Some of them took an interpretive approach and declared that, "*it's not all about the numbers*" [P23]. After 6 months of tracking GPS data, P23 emphasized, "*Numbers are very important, but I think we can aspire for something higher, and I think it also is about the perspective that it allows you to gain*." The general agreement was that important things are found from long-term tracking although people are easily influenced by a day-to-day activity or single data point. The challenge, however, was to find meaningful measures that reflect long-term trends and to keep preserving the initial motivation to tracking even when the latest data point conveys discouraging information.

Tracking Outcomes

Q-Selfers discussed various outcomes of self-tracking, most of which were desirable outcomes such that tracking helped them achieve their initial goals. Many people improved their health and created healthy habits, such as eating healthy, losing weight, and being physically active. Others identified triggers of symptoms and managed to avoid them. P35 realized that driving and drinking coffee were triggers for his panic attacks, and eliminated coffee altogether from his diet, which resulted in a decrease in frequency and severity of the attacks. On the contrary, P8 found a disease that he had not known before, which was ironically a positive outcome for him. P8 initially got into QS to improve his body and get back into shape, but he later discovered that he had Crohn's disease by noticing anomalies from the stool tests he ordered online and his genetic test data. Another positive outcome was the increased awareness of oneself and of the surrounding environment. Being mindful of these things helped people see themselves in a new way such that they were able to understand where ideas came from and how the ideas evolved [P1] or how to create a suitable learning environment for maximizing performance [P41]. Lastly, Q-Selfers reported a positive *reactivity effect* (i.e., change in frequency of the behavior often occurred in the desirable direction) in tracking emotion [P9] and posture [P53]. "*I realized that just by tracking my emotions, I was completely changing them*," said P9, who was able to stabilize her emotional state and prevent herself from experiencing negative emotions.

A few Q-Selfers experienced undesirable outcomes, such as frustration, tracking fatigue, or relapse. Being aware of and confronting negative emotions through tracking caused frustration. "*Because I'm aware of it, it makes it even worse because now I can tell that I'm more anxious than I should be. Before, I was oblivious of being anxious*," said P54, who tracked anxiety and stress. Tracking fatigue was another common outcome of tracking, especially for intensive trackers. After one month of intense tracking on public transit usage, P77 learned many surprising and unexpected findings—such as average commuting time, total cost of using the bus, cost per hour of travel time, and cost per mile. He would not have learned this had he not been tracking, but he said, "*By the end of it, I was really sick of doing it. I just got really fatigued*." However, as the result of a month-long tracking, P77 decided to buy a bike instead of taking the bus because taking the bus was more costly and time-consuming than he had expected. Stopping tracking was not harmful in this case. In fact, P77 used his findings to make a good decision. However, P35, who thought that his panic disorder was under control, stopped tracking and consequently started having panic attacks again. To deal with the relapse and to sustain his commitment to tracking, he built a custom tool to lower the user burden of capturing and looked for a recovery partner for accountability.

Open Challenge: Difficulty in Data Interpretation

Data interpretation was a key hurdle for many Q-Selfers. "*It's not that we lack the information, we're virtually drowning in it. The obstacle is that we don't have the proper tools to interpret the significance of our data,*" said P61, a personal trainer who used to track 11 different things and cross-referenced them with sleep, mood, energy level and acuity. However, after he could not figure out how to extract meaningful information from the 2 years of data, he simplified his tracking strategy to track only two variables. We observed many people who simplified their tracking strategy after their first failed attempt because there was no easy way to analyze and interpret data. Visualizations were helpful in gaining insights, but again, the learning curve for data manipulation (i.e., data cleanup and formatting) and identifying and creating the most appropriate visualization for a given data type was very steep. Helping the general public effectively explore and easily understand their data using visualizations is an active research area for the Information Visualization research community.

IMPLICATIONS FOR SELF-TRACKING TOOL DESIGN

We have identified Q-Selfers' common pitfalls and workarounds when they practice self-tracking. Better self-tracking tool designs could help any potential tracker avoid some of the pitfalls. Here we identify future research efforts that could help address these problems. We discuss how our findings can have broad implications in designing and developing self-tracking technologies.

Provide Early Feedback to Help Identify What to Track

In deciding what to track, Q-Selfers encountered two common pitfalls—tracking too many things, which might cause tracking fatigue, and not tracking triggers and context, which might undermine attempts to gain insights later. When the tracker's motivation is high, it is not a problem to track many things, especially at the beginning of the tracking practice. Tracking multiple things could help people decide which items to keep and which items to stop tracking. What is important then is the self-tracking tool's ability to automate data analysis, provide early feedback on the relationships between different factors, and to suggest eliminating variables that do not seem to correlate with anything. Q-Selfers usually put off data exploration (e.g., running correlations, visualizing data) until later because often, the process involved tedious tasks such as cleaning up data, formatting, and running statistical tests, which could be dramatically reduced by largely automating the process [10]. We envision a self-tracking tool extracting meaningful information, initiating early check-ins, providing real-time visual/textual feedback, and showing comparisons across conditions or correlations among significant factors that can be easily understood.

Support Self-Experimentation by Design

Q-Selfers conducted self-experimentation while compromising scientific rigor. Although innate limitations of self-experimentation (e.g., the experimenter's bias) are hard to avoid, we could help people conduct more rigorous self-experimentation by integrating the single-case research design format [12] into the self-tracking technology design. The three requirements for single-case research design include continuous assessment, baseline assessment, and variability in data [12]. Automated sensing allows easy, unobtrusive, and repeated capturing of a behavior, which could facilitate continuous assessment of the target behavior. Moreover, self-tracking technology could become a platform where people can systematically configure a varying length of baseline and intervention period, tracking frequency, and independent/dependent variables. Quantified-mind [24], a cognitive performance testing website, conveys the similar idea of walking people through setting up self-experimentation. However, cognitive performance is measured by a set of cognitive game scores (hence the fixed dependent variables), which is known to have large practice effects (i.e., repeated testing increases the score).

Maximize the Benefits of Manual Tracking

Sensing and computer automation have many advantages in collecting personal data in terms of reducing mental workload and increasing data accuracy. However, these advantages do not come for free: this automated data collection could reduce awareness and self-reflection resulting from people's engagement with data collection. Q-Selfers expressed that they feel "intimacy with data" when they track data manually. It appears that people make sense of data not only when they explore data but also when they collect data. In addition, some types of data (e.g., subjective sleep quality and pain), by definition, can only be collected via manual tracking. For these reasons, several Q-Selfers built manual tracking tools that drastically lower the user burden, which helped them easily track data and increase awareness. Pushing this idea further, we envision striking a balance between fully automated sensing and manual self-report that can increase awareness, achieve better accuracy, and decrease mental workload.

Promote Self-Reflection

Ironically, the name, Quantified Self is misleading in that it makes people think that Q-Selfers' goal is to quantify their behaviors. It is not. Collecting and quantifying data is just one aspect of QS. The ultimate goal is to reflect upon one's data, extract meaningful insights, and make positive changes, which are the hardest part of QS. HCI research on supporting self-reflection on health monitoring data includes helping people create *unstructured, open-ended* diaries and *sharing* them with others in-person [20]. QS Meetup is another example of engaging people—both speakers and the audience—in self-reflection through storytelling. These are good starting points, and we should further examine ways to support self-reflection on personal data with an aim to enhance *positive reactivity effects*.

CONCLUSION

We analyzed QS Meetup talks and identified that Q-Selfers wanted to improve health, maximize work performance, and find new life experiences through self-tracking. Although many Q-Selfers had positive outcomes from self-tracking, some of them had difficulties throughout the process, such as tracking too many things which led to tracking fatigue, not tracking triggers and context which led to not gaining insights, and lacking scientific rigor which led to inconclusive results. Our goal was to gain insights from Q-Selfers for designing better self-tracking tools in general. Experienced trackers' workarounds might help people avoid some of the pitfalls, but better designs could promote broad adoption of self-tracking technologies by fundamentally boosting their benefits. Specific areas for future research include exploring ways to provide early feedback, to support designing rigorous self-experimentation, to leverage the benefits of—while easing the burden of—manual tracking, and to promote self-reflection. Once a motivated tracker meets a well-designed self-tracking tool, exciting possibilities will arise for gaining insights for health, wellness, and other aspects of life.

ACKNOWLEDGMENTS

We thank Daniel Epstein, Tamara Denning, and A.J. Brush for their feedback. This work was in part supported by Intel Science and Technology Center for Pervasive Computing.

REFERENCES

1. Ali, A.A., Hossain, S.M., Hovsepian, K., Rahman, M.M., Plarre, K., & Kumar, S. mPuff: automated detection of cigarette smoking puffs from respiration measurements. *IPSN '12*, (2012), 269–80.
2. Bentley, F. et al. Health Mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ToCHI,* (2013).
3. Beyer, H. & Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems.* Morgan Kaufmann (1998).
4. Consolvo, S. et al. Activity sensing in the wild: a field trial of Ubifit garden. *CHI '08*, (2008), 1797–806.
5. d3. d3js.org.
6. Fitbit. http://www.fitbit.com.
7. Fox, S., & Duggan, M. Tracking for health. *Pew Internet & American Life Project*, January 28, 2013. http://pewinternet.org/Reports/2013/Tracking-for-Health.aspx, accessed on September 17, 2013.
8. Google Charts. https://developers.google.com/chart.
9. Jawbone UpBand. https://jawbone.com/up.
10. Kandel, S., Paepcke, A., Hellerstein, J., & Heer, J. Wrangler: Interactive visual specification of data transformation scripts. *CHI '11*, (2011), 3363–72.
11. Kay, M. et al. Lullaby: a capture & access system for understanding the sleep environment. *UbiComp '12*, (2012), 226–34.
12. Kazdin, A.E. *Single-case research designs: Methods for clinical and applied settings.* Oxford University Press (2011).
13. Klasnja, P., & Pratt, W. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. *J Biomed Inform 1* (2012), 184–98.
14. Kopp, J. Self-monitoring: A literature review of research and practice. *Social Work Research & Abstracts 24, 4,* Oxford University Press (1988), 8–20.
15. Korotitsch, W.J., & Nelson-Gray, R.O. An overview of self-monitoring research in assessment and treatment. *Psychological Assessment 11(4)* (1999), 415–25.
16. larklife. http://lark.com/products/larklife/experience.
17. Li, I., Dey, A., & Forlizzi, J. A stage-based model of personal informatics systems. *CHI '10*, (2010), 557–66.
18. Li, I., Dey, A., & Forlizzi, J. Understanding my data, myself: supporting self-reflection with Ubicomp technologies. *UbiComp '11*, (2011), 405–14.
19. Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H. Fish'n'Steps: encouraging physical activity with an interactive computer game. *UbiComp '06*, (2006), 261–78.
20. Mamykina, L., Mynatt, E., Davidson, P., Greenblatt, D. MAHI: investigation of social scaffolding for reflective thinking in diabetes management. *CHI '08*, (2008), 477–86.
21. Morris, M., & Guilak, F. Mobile heart health: project highlight. *Pervasive Computing 8(2)* (2009), 57–61.
22. Nike Fuelband. http://www.nike.com/cdp/fuelband.
23. Patel, R.A., Klasnja, P., Hartzler, A., Unruh, K.T., & Pratt, W. Probing the benefits of real-time tracking during cancer care. *AMIA '12 Fall Symposium*, (2012).
24. Quantified Mind. http://www.quantified-mind.com.
25. Quantified Self. http://quantifiedself.com.
26. Rosenthal, R. *Experimenter effects in behavioral research*. Appleton-Century-Crofts (1966).
27. Swan, M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. *Int J Environ Res Public Health 6(2)* (2009), 492–525.
28. Technology Review. http://www.technologyreview.com/news/514356/stephen-wolfram-on-personal-analytics.
29. Troshynski, E., Lee, C., & Dourish, P. Accountabilities of presence: reframing location-based systems. *CHI '08,* (2008), 487–96.
30. Wolfram. http://www.wolfram.com.
31. Zeo. http://myzeo.com.