Reconsidering the Device in the Drawer: Lapses as a Design Opportunity in Personal Informatics

Daniel A. Epstein[1], Jennifer H. Kang[1,2], Laura R. Pina[1,3], James Fogarty[1], Sean A. Munson[3]
[1]Computer Science & Engineering [2]The Information School [3]Human Centered Design & Engineering
DUB Group, University of Washington
{depstein, jkang331, lpina, jfogarty}@cs.washington.edu, smunson@uw.edu

ABSTRACT
People stop using personal tracking tools over time, referred to as the *lapsing* stage of their tool use. We explore how designs can support people when they lapse in tracking, considering how to design data representations for a person who lapses in Fitbit use. Through a survey of 141 people who had lapsed in using Fitbit, we identified three use patterns and four perspectives on tracking. Participants then viewed seven visual representations of their Fitbit data and seven approaches to framing this data. Participant Fitbit use and perspective on tracking influenced their preference, which we surface in a series of contrasts. Specifically, our findings guide selecting appropriate aggregations from Fitbit use (e.g., aggregate more when someone has less data), choosing an appropriate framing technique from tracking perspective (e.g., ensure framing aligns with how the person feels about tracking), and creating appropriate social comparisons (e.g., portray the person positively compared to peers). We conclude by discussing how these contrasts suggest new designs and opportunities in other tracking domains.

Author Keywords
Personal Informatics; Lapsing; Abandonment; Re-engagement.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g., HCI).

INTRODUCTION
Personal informatics, defined as the process of collecting and reflecting on personal data [29], is now common practice [17]. However, people often have trouble maintaining tracking, for example keeping their devices charged [13,22] or forgetting to wear them [13]. Epstein et al. refer to this idea that people stop actively tracking as *lapsing* [13]. Some people return to tracking, either by resuming use of the same tool or selecting a new tool that better serves their needs or is easier to maintain using. Others *abandon* tracking entirely. Although some abandon tracking because it provided little value [11], others leave tracking with feelings of frustration that tracking did not help them achieve their goals [11,27] or guilty they could not sustain the habit of tracking [7,11].

UbiComp '16, September 12-16, 2016, Heidelberg, Germany

http://dx.doi.org/10.1145/2971648.2971656

Despite research finding that many people lapse or abandon tracking and that tracking goals are often left unmet, there is little advice on how to support people after a tracking lapse. We posit that inviting people who have lapsed to review data they previously collected can offer support and encouragement, particularly if the data is surfaced in new or interesting ways. A positive outcome in this case may or may not be a resumption of tracking. For example, a person may instead benefit from a reminder of the insights they gained from tracking.

This work begins to explore the broad space of potential designs that can support people after they lapse in tracking. We limit this initial exploration to designing for people who have lapsed in Fitbit use. Fitbit [15] is a digital pedometer and is currently the largest seller of wearable technology, with over 21 million units sold in 2015 [25]. Physical activity is a well-studied and understood domain of self-tracking in the research community [7,10,13,27]. Fitbit specifically is often used in studies of physical activity tracking [5,28,35].

We further limit our exploration to familiar visualizations of activity data (e.g., bar charts) with natural language summaries. Prior work has shown these can help people learn from their personal data [2,12,24]. Although other designs have been studied (e.g. abstract displays of activity [8,30]), simple visualizations and natural language summaries are more common in both research prototypes and commercial tools.

We specifically consider two aspects of data presentation: the selection of a *visual cut* and the use of a *framing technique*. The visual cuts technique suggests that distilling a person's complex multi-dimensional personal informatics data into simple visualizations can yield interesting, understandable, and actionable findings [12]. The choice of framing technique has been shown to affect self-efficacy in the context of Fitbit data [5]. Framing feedback in neutral or positive ways (e.g., highlighting achievement and not punishing inactivity) has been proposed as a design principle for behavior change technologies [8]. Persuasive technology literature suggests additional strategies, such as asking a person to identify barriers or explicitly promoting a return to tracking [32].

We surveyed 141 people who had lapsed in their use of Fitbit, recruited through a combination of Amazon Mechanical Turk and a snowball sample. We used a mixed-methods approach to categorize them into three patterns of use and four perspectives on their Fitbit experience. The participants reviewed their own Fitbit data presented in seven visual cuts, randomly paired with seven framing techniques.

We report on findings from these two analyses, including:

- Identifying three Fitbit use patterns: short use of less than four months, long and consistent use, and intermittent use.
- Preferences for feedback correlated with how long people had used their Fitbits. Participants who used their Fitbits for less time preferred aggregated views, such as by hour or by day of the week. Participants who had used their Fitbits longer had the opposite preference, valuing timelines and histograms of their use, aggregated by month or year.
- Participants who felt guilt or frustration over abandoning Fitbit tracking appreciated messages encouraging them to return. However, participants who thought tracking was a useless or negative experience felt these messages were presumptive and inconsistent with their experiences.
- Social comparisons can be encouraging and provide interesting information, even to participants who thought tracking was useless or negative. However, care needs to be taken to ensure the comparisons are desired and reflect positively on the person tracking.

We conclude by discussing how contrasts in responses to visualizations suggest new designs as well as opportunities for other common types of personal informatics data.

RELATED WORK

Prior personal informatics research has considered why people lapse in their use of tracking tools, how to present collected data, and how to frame messages in ubiquitous systems.

Lapses and Abandonment of Personal Informatics Tools

In previous work, we draw a distinction between people who have temporarily *lapsed* in tracking versus people who *abandon* tracking altogether [13]. People may lapse for a variety of reasons, including having difficulty remembering to keep their devices charged [13,22] or forgetting to wear them [13]. These challenges can lead to incomplete records, which in turn often lead people to abandon the habit. For example, people who track their food consumption often struggle to record everything they eat. These lapses result in incomplete data, which diminishes the utility of a calorie budget, and can ultimately lead people to abandon the practice [9].

Other people abandon tracking without a substantial lapse. People abandon tools because collected data is not useful [27] or they misjudged their dedication to tracking [7]. Many feel guilty for abandoning tracking or frustrated that tracking was unsuccessful at changing behaviors [11]. Others become "obsessed" with self-quantification, so abandoning tracking leaves them feeling liberated from this burden [11]. Finally, many abandon after successfully accomplishing their goal, such as losing weight or understanding their habits [7,11].

People do not necessarily want to return to tracking after abandonment. Although many people express an intention to return to tracking [27], people who accomplished their goal for tracking or found tracking unpleasant often do not want to return [11]. To this end, Clawson et al. develop the concept of "happy abandonment", suggesting that tools should support people who no longer want to track [7].

Understanding Personal Informatics Data

Methods for helping people make sense of personal data vary widely. Although particularly skilled and motivated people may make sense of their own data through visual or statistical analysis [6], most rely on presentations created by others.

Visualizations [12,24], natural language sentences [2,12,21], and abstract presentations [8,14,30] can help people understand their habits. Prior work has expressed the importance of highlighting data and concepts important to the person tracking, with specific data and concepts varying from person to person [12]. Results should also highlight findings that are interesting (i.e., findings that present the viewer with novel information or that address or trigger some curiosity). In contrast, some participants in the Health Mashups system described the correlations surfaced between tracked data as "obvious," which hindered the usefulness of the system [2].

Framing in Self-Tracking

Strategies developed in the fields of persuasive technology [1,16] and behavior change [31,32] are often employed in the design of technology to support increased activity. Such work has developed a set of strategies for persuading and supporting behavior change, which we also utilize in this work.

Designers must consider not only *what* information they present to a tracker, but also design *how* they frame that information. Fish'n'Steps presented low step totals in the form of a sad fish, which discouraged inactive participants from looking at their data [30]. Based on this finding, Consolvo et al. designed UbiFit Garden to only highlight achievement. Low activity was therefore represented by an empty garden [8]. Further, Choe et al. found participant self-efficacy was higher when steps walked were framed as achievement (e.g., how many steps you walked) rather than a shortcoming (e.g., how many steps you have left to reach your goal) [5].

These findings suggest positive framing of activity promotes self-efficacy and engagement. Additional strategies that have been suggested as viable include provocative messages [1] and prompting people to consider potential barriers [32]. For example, the Habito app presented participants with both informational (e.g., "*you've burned 200 calories by walking today*") and persuasive (e.g., "*try walking more when you're on the phone*") messages [21]. Upon receiving persuasive messages, participants took longer to re-engage with the app, but were more likely to both start walking and walk further.

PRESENTING FITBIT DATA TO LAPSED TRACKERS

We now consider design opportunities for people who have lapsed in their practice of self-tracking. We specifically evaluate designs for physical activity tracking, possibly the most-studied domain of self-tracking and one that is often studied in making broader claims about personal informatics [13,19,33]. We specifically look at Fitbit [15], the largest seller of wearable technology [25]. Considering only one tracking platform is a limitation of this work, but it allowed us to limit development overhead and design for a particular set of constraints. We will later discuss how to extend these design opportunities to people tracking other types of data.

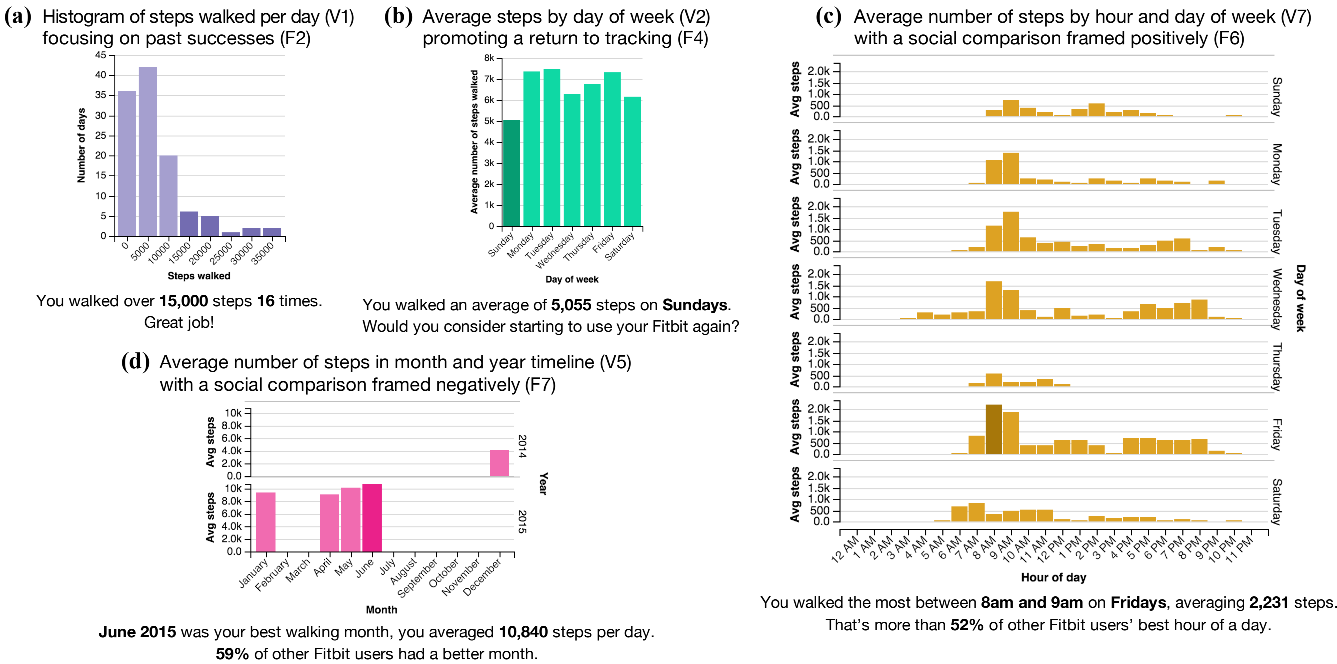


Figure 1. We explored how people who have lapsed in their use of Fitbit responded to representations of their data. We varied the visual cut presented (e.g., aggregated by day, month) and how the data was framed (e.g., social comparisons, suggesting resumption). The full set of presented cuts and framings are in Table 3 and discussed in our analysis.

Each representation is defined by two components: a visual cut and a framing technique. Figure 1 presents four examples from our study, each pairing a visual cut with a framing technique. Table 3 in our analysis enumerates the full set of visual cuts and framing techniques examined in this work.

Data Collection

HCI researchers have noted that engagement measures such as self-efficacy and knowledge can provide deeper design insight than studies focused only on outcome measures, such as attempting to measure behavior change [26]. Following this guidance, we conducted an online experiment focused on intermediate measures, such as whether messages felt judgmental or appropriate and how much the recipient believed they could succeed at being physically active.

We first conducted a pilot survey on Amazon Mechanical Turk (AMT) with 33 participants from the United States. We then refined the survey questions and recruited an additional 141 participants from both AMT and a snowball sample recruited through posts to social media and university mailing lists. All participants were required to have worn their Fitbit at least three days but not tracked steps in the past week. They further acknowledged that they were not using another app or device to track their physical activity. To ensure response quality and limit spam, we restricted AMT participation to workers with at least a 95% HIT acceptance rate and 1,000 completed HITs. AMT workers were compensated $2.00 for completing the survey, and snowball sample participants were entered into a drawing for one $100 and one $50 Amazon gift card. This resulted in similar expected value of compensation, provided in the form appropriate for the norms of each recruitment channel.

The survey included two major parts. Participants first (1) logged into Fitbit, granted access to their data via the Fitbit API, and then answered a few questions about their experiences using Fitbit. They then (2) were presented with seven representations of their own Fitbit data, provided their opinions on each, and gave overall thoughts on viewing their Fitbit data. Finally, they completed a short demographic survey. The snowball sample included an optional final step where participants could see their Fitbit friends who had also lapsed. Participants who agreed to view this screen were asked to forward the survey to their friends who had lapsed.

190 people completed Part 1 of the survey, of which 141 qualified for our study (i.e., had lapsed in Fitbit use). Of the 49 excluded participants, 35 had used a Fitbit within the past week, 1 used a Fitbit for only a single day, and 13 indicated in qualitative responses they had switched to another tracking tool (e.g., "*I have an Apple Watch now*"). We omit the 13 because we wanted participants who had lapsed in their tracking (i.e., not switched to tracking with another tool). Our analyses therefore report on 141 participants.

Each representation in Part 2 presented the participant's Fitbit data in a different visual cut implemented with Vega-Lite [38]. Participants rated each representation on five 7-point Likert scales described in our analysis section. Finally, participants provided open-ended reactions in a free response text field. 124 of the participants saw all 7 representations (average 6.6) and then completed the final demographics section (Table 1). We thus report on 915 participant reactions to a visualization.

We conducted two analyses of data, corresponding to the two major parts of the survey. We first sought to characterize participant tracking experiences, which we believed would

	AMT Fitbit Use (74 people)	Snowball Fitbit Use (67 people)
Days Tracked	avg 151, stdev 132 min 4, max 603	avg 234, stdev 212 min 8, max 1053
Days Between First and Last	avg 295, stdev 251 min 3, max 1035	avg 372, stdev 290 min 16, max 1074
Average # Steps	avg 6520, stdev 2835 min 1548, max 17657	avg 7931, stdev 2434 min 4053, max 16459
Days Since Used	avg 224, stdev 264 min 7, max 957	avg 280, stdev 260 min 13, max 991
	AMT Demographics (71 people)	**Snowball Demographics (53 people)**
Gender	51 female, 20 male	33 female, 19 male 1 no response
Age	avg 33.6 min 22, max 59	avg 30.9 min 18, max 50

Table 1. We recruited people who lapsed in Fitbit use from Amazon Mechanical Turk and a Snowball Sample.

inform which visualizations and framing techniques they found most appropriate. We then analyzed the intervention data in light of these participant tracking characterizations. It is important to note these analyses were conducted in serial to not bias the results. We first characterized participant tracking patterns based on their Fitbit data and responses in Part 1, then analyzed the impact of the intervention undertaken in Part 2.

Limitations

Relative to participants recruited through the snowball sample, participants recruited on AMT wore their Fitbits for fewer days (t_{109}=2.74, *p*=0.007, 95% CI 23-143 fewer days) and averaged fewer steps per day (t_{139}=3.16, *p*=0.002, 95% CI 527-2295 fewer steps). We cannot know if either sample is representative of people who have lapsed in Fitbit use, but believe the sample is sufficiently large and diverse to offer valuable perspective.

Prior work suggests 30 days without tracking is a reasonable indication a person has lapsed [13,35]. We selected 7 days for recruitment because the Fitbit API only allowed access to friend step total for the past week, so a 30-day cutoff was impractical for snowball recruiting. 24 participants stopped tracking between 7 and 30 days before completing the survey. We did not find any differences between these 24 participants and the 117 who lapsed more than 30 days before the study. In Part 1 of our analysis, the use patterns and perspectives of these participants were similar to participants who had lapsed for more than 30 days. We also conducted a robustness check on Part 2 of our analysis by removing these participants. Our results were similar, so we therefore include these participants.

Finally, we restricted AMT participation to the United States, and we expect nearly all of our snowball sample was also from the U.S. Other cultures may respond differently to the framing techniques we used. For example, Hofstede's collectivism cultural dimension differentiates group ties in a society, with the U.S. as a less collectivist society [23]. In contrast to framings that compare performance between people, more collectivist societies may prefer framings that focus on grouping people's performance together.

PART 1: FITBIT USE PATTERNS AND PERSPECTIVES

We used a mixed-method approach to categorize participant Fitbit use. Three researchers created an affinity diagram of participant data collected in Part 1 of the survey, holistically considering each participant's responses to all the questions. We defined use patterns from the pilot study with 33 participants, then applied the use pattern definitions to the larger pool of participants. Using the patterns found during affinity diagramming, the lead researcher then applied an Affinity Propagation clustering algorithm to the data [18].

We originally hypothesized there would be a strong connection between use patterns and perspective on tracking. However, it became clear that participant use patterns were only loosely correlated with their perspectives on tracking. We therefore separately characterized use patterns and perspectives. Three codes for use pattern and four for perspective emerged.

Initial agreement between the three researchers coding the data was high for use pattern (Fleiss's κ=0.78) and moderate for perspective (Fleiss's κ=0.51). Researchers resolved final categorizations through arbitration (summarized in Table 2).

Three Use Patterns

We grouped participants into three patterns of Fitbit use: short use, long and consistent use, and intermittent use. These themes emerged from considering three variables:

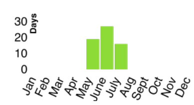
Days Tracked: we define a *day tracked* as any day with at least one step tracked. This may slightly over-count by including days with very few steps, but serves as an approximate measure for Fitbit use that is independent of activity level.

Trials: for analysis, we define a *trial* as a period of Fitbit use separated by at least 30 days. People do not necessarily use their Fitbit every day during a trial (e.g., forgetting, opting not to track a day). However, prior work has suggested 30 days to be a reasonable indication of a person who has lapsed [13,35].

Consistency of Tracking: we define *consistency* as the percentage of days tracked relative to the number of days during a trial. A ratio of 100% thus indicates a person who used their Fitbit every day during a trial.

Short Use

We categorized *short use* as four months or fewer with steps recorded. We found the four month cutoff created greater separation from other patterns of use, and it is consistent with a few-month use pattern identified in prior work [27,35]. 89.8% of short use participants tracked during a single trial. Tracking by this group was often inconsistent, with participants recording steps an average of 76.9% of days during a trial.

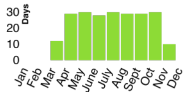


p131's Fitbit Use

More than other use patterns, participants with short use stopped because they lost interest: "*I enjoyed it, but my interest level petered off*" (p28). This group also more often got their Fitbit from friends, significant others, or corporate wellness programs: "*wife uses hers, and she got me one as a gift*" (p117).

Long and Consistent Use

Participants with *long and consistent use* had at least five months with any step data. Similar to the short use pattern, 78.6% of participants in this group tracked during a single trial. The remaining participants used their Fitbit for two trials that were separated by a month or two. Participants with long and consistent use tracked more consistently than participants with short use (t_{138}=2.79, p=0.017, 95% CI 1.5-18.2% more consistent), recording steps an average of 86.7% of days during a trial.

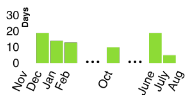


p46's Fitbit Use

Participants with long and consistent use were more likely to stop because devices were lost or broken. p144 considered replacing his broken Fitbit, but "*then I found I did not miss it so I did not replace it.*" Conversely, p50 would track again but cannot afford a new Fitbit: "*I loved it. I only stopped using it when I lost it, and I can't afford to replace it.*"

Intermittent Use

We categorized use as *intermittent* if participants tracked in three or more trials. These participants tracked less consistently than participants with long and consistent use (t_{138}=-3.27, p=0.004, 95% CI 3.2-19.8% less consistent) with no detectible difference in consistency versus participants with short use (p=0.877). They tended to have more days with steps recorded than participants with short use (t_{138}=5.42, p<0.001, 95% CI 90-230 more days) and fewer days with steps recorded than participants with long and consistent use (t_{138}=-3.25, p=0.004, 95% CI 27-173 fewer days).



p92's Fitbit Use

Participants with intermittent use identified barriers to tracking that caused them to track intermittently. p22 started tracking to receive discounts on her health insurance, but "*somewhat resented the daily requirement to wear it.*" Charging the Fitbit was too burdensome for p85: "*my motivation eventually ran out due to this tedious nature of having to charge the device.*" p124 described their Fitbit use pattern as cyclic: "*I feel like I easily start and stop - I really like it, then get bored with it, then like it again, etc.*"

Four Perspectives on Tracking Experience

Participants looked back on their tracking experience with different perspectives. The four perspectives we identify in this work draw heavily from prior work [11], so we only briefly explain them here. Our current work goes beyond that prior examination by considering willingness to resume using a Fitbit as well as relationships between these perspectives and participant responses to representations of their tracking data.

Learned Habits

Five participants found "*I learned a lot from my Fitbit experience and no longer need it to keep track of me*" (p144). p93 said his "*curiosity about daily activity levels is satisfied,*" and thus felt no reason to continue tracking. Participants with this perspective were generally lukewarm or negative on resuming.

That few participants had this *learned habits* perspective is both consistent with prior work [11] and an indicator that people who used Fitbit are generally not in line with the notion of "happy abandonment" [7]. The small portion of our sample with this perspective limits our analysis, but the distinctiveness of this perspective made it inappropriate to combine this group with another group for analysis.

	N	Days Tracked		# Trials		Consistency	
Short Use	49	avg	57	med	1	avg	77%
		min	4	min	1	min	17%
		max	128	max	2	max	100%
Long and Consistent Use	42	avg	316	med	1	avg	87%
		min	113	min	1	min	57%
		max	1053	max	2	max	100%
Intermittent Use	50	avg	216	med	3	avg	75%
		min	8	min	2	min	38%
		max	621	max	6	max	100%

	N	Would Return		Representative Perspectives
Learned Habits	5	Agree:	20%	*"I feel that I know my habits enough now"*
		Neutral:	40%	
		Disagree:	40%	
Useless or Negative	21	Agree:	24%	*"It didn't do anything useful for me"*
		Neutral:	0%	*"Don't want to fall into obsessive habits again"*
		Disagree:	76%	
Guilt or Frustration	70	Agree:	97%	*"Although I have gotten off track, I'd like to be more conscientious about using my Fitbit… it really did help me."*
		Neutral:	3%	
		Disagree:	0%	
Conflicting Feelings	45	Agree:	76%	*"I have one, I could be getting data, but I'm lazy and don't"*
		Neutral:	18%	*"I liked using my Fitbit, but I don't have the money right now for a new one"*
		Disagree:	7%	

Table 2. Participants were almost evenly split by use pattern. Few participants felt like they had learned their habits, and nearly half felt guilt or frustration over abandoning tracking.

Useless or Negative

Some participants felt they did not get value from tracking, finding it *useless or negative*: "*I didn't really get anything from it, and it was a (very minor) inconvenience*" (p104). p34 struggled to connect tracked data to changing his behavior "*I thought it gave me interesting information, but didn't change my habits at all.*" Two participants further expressed concern tracking was encouraging obsessive practices. P14 "*got a bit too obsessive about getting my steps*" and did not want to return to Fitbit out of fear of resuming those habits. Based on their experiences using Fitbit, most participants with this perspective stated they would not use a Fitbit again.

Guilt or Frustration

Many felt either guilty that they were not using their Fitbit or frustrated that Fitbit did not help them change their behavior. These participants tended find Fitbit valuable, such as p60: "*it was a useful tool to weight loss,*" and p107: "*it made me take extra steps in a day.*" They often blamed themselves for abandoning Fitbit, such as p91: "*I shouldn't give up, but rather try again differently and hope to get different results.*"

Participants with this *guilt or frustration* perspective were interested in using a Fitbit again. Six participants were curious how their habits had changed since lapsing: "*I don't have the same lifestyle I used to have when I wore a Fitbit, so I want to see how poorly or better I'm doing compared to that time*" (p97). Others found Fitbit motivating, and wanted that motivation again: "*Fitbit motivated me to walk more… I feel a bit empty without it*" (p51). About half of these participants stopped tracking because they lost their Fitbit or did not replace it after it broke (including p51). The remainder could not motivate themselves or remember to use Fitbit again

Visual Cut	Randomly ordered and paired	Framing Technique
V1: Histogram of number of steps walked each day		F1: Prompt consideration of barriers
V2: Average number of steps by day of week		F2: Focus on past successes
V3: Timeline of average number of steps by month		F3: Setting an improvement goal
V4: Timeline of number of steps by month and year		F4: Promote a return to tracking
V5: Timeline of days using Fitbit by month and year		F5: Suggest person does not need need to track
V6: Average number of steps by hour		F6: Social comparison framed positively
V7: Average number of steps by hour and day of week		F7: Social comparison framed negatively

Table 3. Participants saw 7 visual cuts through their own tracked data, each presented with a different framing technique.

"*I would like to give it another chance. I just kept forgetting about it*" (p12). p110 described both a health barrier and a lack of motivation stopping her from returning: "*I need to become more motivated and exercise more. However, I have health conditions holding me back and lack of motivation.*"

Conflicting Feelings

Some participants expressed *conflicting feelings* in describing their perspective on using Fitbit, finding tracking both valuable and a difficult habit to maintain. These participants did not cleanly fit in the above groups. They typically did not express guilt or frustration over lapsing, but rather ambivalence: "*I do not feel strongly either way. I guess if I had a good reason [to return to Fitbit] I would*" (p145).

Participants in this group felt less sure they wanted to track, despite seeing benefits and learning about themselves: "*I want to be healthier and I'd like it to work, but I'm not sure it's the best method for me*" (p2). p63 found tracking useful, but was "*not ready to commit to wearing it everyday, again.*"

Many in this group described positive experiences with Fitbit, but stopped for reasons unrelated to tracking (e.g., p39: "*injury caused me to be able to walk less*") or because the device did not support their use case (e.g., p66: "*Fitbit couldn't handle me being on a bike very well*"). Four participants described wanting to try newer Fitbits, including p137: "*I would love to try the newer models… to see if they have improved.*"

PART 2: REACTIONS TO REPRESENTATIONS

We now consider how Fitbit use patterns and perspectives influence participant preferences for data representation. Each representation consisted of a *visual cut* paired with a *framing technique*. Table 3 summarizes each of these.

Visual Cuts: We intended to promote self-understanding by highlighting interesting aspects of the Fitbit wearer's tracking experience. People typically have varied preferences for what information is surfaced, so we consider a variety of potential cuts [12].

The cuts we selected were informed by data aggregations explored in prior work. We separated cuts into three groups: a *histogram* view (V1), views finely *aggregated* (V2, V6, V7), and *timeline* views (V3, V4, V5). V1's histogram view (example in Figure 1a) was intended to identify days of particularly high or low activity [12]. V2 and V7 (Figure 1b, 1c) are based on Bentley et al.'s approach of identifying correlations between streams of self-collected data that is aggregated by day of the week [2]. V6 and V7 (Figure 1c) are based in fine-grained summaries of physical activity data, as previously considered in the context of social sharing [10]. V3, V4, and V7 (Figure 1d) are based in aggregations by month and year, inspired by aggregations of electrical and water consumption data [20].

Visualizations can be grammatically composed through *marks* (e.g., wedges of a pie chart, bars of a bar chart), *visual properties* (e.g., position and color), and *data* [4]. Because the data presented varied between visual cut, we chose to fix other visualization properties as much as possible. As such, we presented all cuts with a common mark: bar graphs. This said, some participants wanted other representations. p72 said, "*I think it would be more interesting as a line graph with the years each being a different color.*" We varied color for visual differentiation of the different cuts. For simplicity, we did not attempt to visualize data spread. We did not include error bars and every bar represents a count or average.

Framing Techniques: The presented framing techniques drew from taxonomies and strategies in persuasive technology and behavior change that could be logically mapped to Fitbit data [1,16,31,32]. Behavior change and persuasive technology literature suggest other strategies that we did not evaluate, such as providing information on when and where to perform a behavior [32]. We selected strategies we felt could be derived from only Fitbit data (i.e., only steps and times). Suggesting when and where to perform a behavior would require more contextual information than can be inferred from step data and may come across as judgmental. Considering other framing techniques is an interesting opportunity for future work, particularly with other data types (e.g., location, heart rate).

Depending on the framing technique, one or more bars were highlighted to convey the framing. Highlighted bars were summarized, and a statement was included that was intended to be either informative or persuasive.

Our framing technologies included three pairs that offered opposing perspectives. F1 suggested the participant consider what prevents them from walking more, while F2 (Figure 1a) focused on successful activity periods [32]. F4 (Figure 1b) and F5 offered differing views on whether the participant should try to use their Fitbit again [16,32]. F6 and F7 provided social comparisons framed differently, a common strategy in technology to support behavior change [19,20,37]. These representations highlighted the percentage of Fitbit wearers the individual outperformed (F6, Figure 1c) versus the percentage who outperformed the individual (F7, Figure 1d). Percentages were generated by comparing participants against an existing dataset of 114 people who had used Fitbit for at least 3 months, collected for a different ongoing study.

Methods of Analysis

Our analyses seek to address the following design questions:

1. How should an individual's Fitbit use pattern influence the selection of a visual cut used in presentation?
2. How should an individual's perspective on tracking influence the framing technique used in presentation?
3. How should social comparisons be used in presentation?

We analyze participant opinions with respect to different aspects of visual cuts and framing techniques. In Part 2 of our study, participants responded to five 7-item Likert scales for each representation. They rated agreement or disagreement with statements of whether they thought a representation was (a) appropriate, (b) informative, (c) judgmental, (d) made them feel like they could succeed at being healthy, and (e) made them feel like they could succeed at returning to tracking.

We analyzed our data using a series of generalized mixed-effect models with the Likert data as an ordinal response. We include demographic information, recruitment method, patterns of Fitbit use, and number of representations seen as fixed effects. Participant ID was included as a random effect to account for individual differences. Visual cut and framing techniques were treated as categorical fixed effects. We found no significant effects of presentation order, age, gender, activity level, or time since stopping tracking. We therefore do not report on these factors further. Compared to participants recruited through the snowball sample, AMT participants rated representations more appropriate (Z=2.49, *p*=0.013, 95% CI 0.2-1.4 higher on a 7-point Likert scale), less judgmental (Z=-2.10, *p*=0.036, 95% CI 0.1-1.4 lower), and found representations made them feel more like they could succeed at returning to tracking (Z=2.18, *p*=0.029, 95% CI 0.1-1.6 higher). We suspect AMT payment incentivized more favorable ratings. We include recruitment method in our model to account for this, but do not further report on it as it does not pertain to our research questions.

Our study design was fully factorial, but randomly assigned. All visual cuts were paired with all framing techniques and seen by at least 7 participants (max 26). Given the small N and relative unevenness of each of the 49 conditions, our analysis focuses on the main effects of visual cut and framing technique. Interaction effects may exist (e.g., certain framing techniques which pair well with certain visual cuts), but it is first important to consider broad recommendations in this new design space before seeking an "ideal" pairing.

We first explored the data with Tableau [36] to understand the patterns and trends. From this exploration, it became clear that use pattern had a substantial effect on cut preference and that tracking perspective had a substantial effect on framing technique preference. We thus considered interaction effects between use pattern and visual cut and between tracking perspective and framing technique.

Mixed-model analyses typically encode one level of a factor as a "reference" level to which all other levels are compared. In this study, we are interested in comparing across levels (e.g., across cuts or framings), not against a "reference."

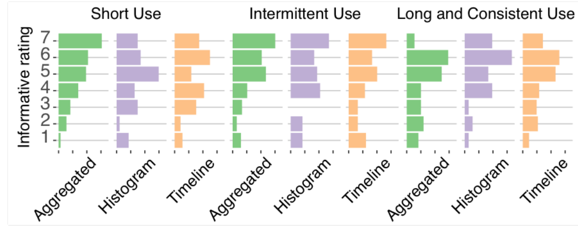


Figure 2. Informative rating (7 indicates strong agreement) considered by visual cut and use pattern. Participants with *short* or *intermittent* use preferred aggregated representations. Those with *long and consistent* use preferred histograms and timelines.

The results presented in this paper are post-hoc pairwise and contrast tests on our models, with a Bonferroni correction applied to account for multiple comparisons.

Selecting Appropriate Aggregations Based on Fitbit Use

Contrast 1: Visual cut preference varied based on use pattern. Participants with less collected data preferred cuts aggregated by hour or day, while participants with more data preferred cuts which highlighted their long use.

Participants with *long and consistent use* rated aggregated cuts as less informative than participants with the other two use patterns (Z=-3.54, *p*=0.001, 95% CI 0.4-1.5 decrease on a 7-point Likert scale). Moreover, these participants found aggregated cuts less informative than histograms (Z=-3.02, *p*=0.007, 95% CI 0.2-1.9 decrease). Participants with *short use* of Fitbit (less than 4 months) expressed the opposite preference, finding aggregated representations more informative than both histograms and timelines (Z=3.79, *p*<0.001, 95% CI 0.2-0.9 increase). Figure 2 summarizes this data with *informative* rating as the response.

Participant comments offer some explanation. Participants with longer Fitbit use tended to have a better understanding of their daily habits, and thus did not find the aggregations as informative. p113 said she was "*not surprised with when my activity takes place,*" a sentiment shared by 12 others. p101 felt the cut "*needs to highlight days lapsed (rather than just avg across weeks) to show where problem areas are.*"

Participants with *long and consistent use* were more excited to see a cut that highlighted their long use pattern. p44 felt "*seeing how many days I succeeded in wearing [my Fitbit] is motivating and inspiring to me.*" By contrast, this information can be discouraging to people with little data. p53 used her Fitbit for 43 days and felt "*I'm disappointed in myself for not using the Fitbit for longer and giving up.*"

Participant preferences thus suggest visual representations should align with the amount of data collected. A person who tracked for a long amount of time may implicitly understand their daily or weekly habits (e.g., reflection in action [34]). A person with less tracking experience or an intermittent habit may be more likely to benefit from seeing information which reveals their habits through aggregation.

Choosing Framing Technique from Tracking Perspective

Contrast 2: Participants preferred positive and encouraging framings to critical or introspective framings.

We contrasted framing techniques that asked participants to consider their barriers to being active (F1) with techniques focusing on past successes (F2). Independent of tracking perspective, participants rated F1 less informative (Z=-2.02, p=0.043, 95% CI 0.02-1.5 decrease on a 7-point Likert scale), less appropriate (Z=-2.16, p=0.031, 95% CI 0.1-1.4 decrease), and more judgmental (Z=3.66, p<0.001, 95% CI 0.8-2.7 increase) than F2. Participants felt being asked to consider barriers was "*snarky*" (p56, p147), "*not motivational*" (p65, 3 others), and "*judgmental*" (p126, 7 others). By contrast, participants found focusing on successes "*encouraging*" (p7, 3 others) and "*positive*" (p107, 16 others). This is consistent with prior work suggesting neutral-to-positive framing [5,8].

Contrast 3: The more interested a participant was in returning to tracking, the more they felt viewing representations of their prior tracking data encouraged them to do so.

Participants who felt *guilt or frustration* that they abandoned tracking were more likely to agree that they could succeed at returning to tracking than all other groups (Z=4.47, p<0.001, 95% CI 0.7-2.6 higher on a 7-point Likert scale). Participants with *conflicting emotions* agreed less than participants who felt *guilt or frustration* (Z=-2.60, p=0.046, 95% CI 0.1-2.2 lower). Results were inconclusive about whether they agreed more than participants who thought tracking was *useless or negative* (p=0.098, 95% CI -0.2-3.0).

These differences validate division of participant perspectives. We would expect participants who felt *guilt or frustration* would rate themselves more willing to return to Fitbit, while those who felt tracking was *useless or negative* would be less willing. As one might expect, participants with conflicting feelings rated themselves between those two groups.

Contrast 4: Participants preferred framing techniques aligned with their perspective on whether they should return to tracking.

We contrasted two framing techniques offering opposing perspectives on whether the participant should return to tracking, highlighted in Figure 3. F4 promoted resumption, while F5 suggested the participant did not need to track anymore. Participants who felt *guilt or frustration* rated F4 made them feel more like they could return (Z=2.34, p=0.019, 95% CI 0.1-1.5 higher) and more informative (Z=2.04, p=0.041, 95% CI 0.03-1.3 higher) than F5. Shown a representation framed with F4, p96 said "*this data proves to me I can do this… just have to get started again*", a sentiment shared by 8 participants. In contrast, shown a representation framed with F5, p64 said "*I feel like I really need a Fitbit, so I don't really trust this statement*" (16 others agreed).

Conversely, participants who thought tracking *was useless or negative* rated F4 less informative (Z=-2.34, p=0.019, 95% CI 0.23-2.62 lower) than F5. p61 found F5 "*odd and somewhat positive.*" p34 found the suggestion a bit unusual: "*seems like it isn't a good marketing idea.*" p104 stressed, "*I never *needed* it. All these quotes assume I'm trying to get more active, and I'm not.*" These participants reacted even more negatively to F4, including p47 "*I don't like being solicited to do something I consciously chose to stop doing because I felt it was no longer beneficial*", a sentiment shared by 4 others.

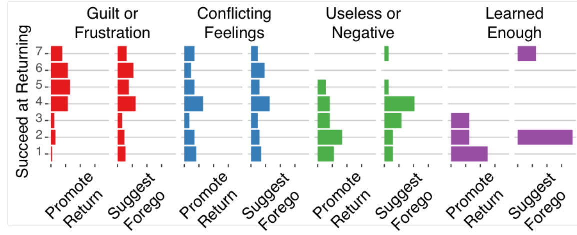


Figure 3. Participants who felt *guilt or frustration* preferred framing techniques which suggested they return to tracking. Participants with *conflicting feelings* were more mixed.

Participants with *conflicting feelings* did not demonstrate a strong preference toward F4 or F5. We found no detectable difference for these participants for feeling like they could return (p=0.201), inconclusive results for whether F4 was less informative than F5 (p=0.056, 95% CI -0.02-1.54). Qualitative responses were mixed. p32 and 7 others found the recommendation to return "*made me reconsider using my Fitbit again and that felt good.*" p150 and 5 others found the suggestion antagonizing: "*So, I put the Fitbit away for 3 months. So what. If I want to track, I can. This is very negative for me. Very demotivating.*"

These results suggest framing approaches should align with the perspectives of the person viewing the data. As prior work has suggested, people do not necessarily want to track after lapsing [7,11,13,27]. Designers should be sensitive to the perspective of people who do not wish to return to tracking. However, messages suggesting a return can be beneficial to trackers interested in returning, such as p92: "*[this message] helps to see how much I am walking and encourages me to start using my Fitbit again.*"

Creating Appropriate Social Comparisons

Contrast 5: Participants preferred social comparisons framed positively over those framed negatively, even when the comparisons represented the same information.

We contrasted the framing techniques which framed social comparisons positively (F6) and negatively (F7). Participants across use pattern and perspective rated F6 more appropriate (Z=2.62, p=0.009, 95% CI 0.3-1.7 higher on a 7-point Likert scale), less judgmental (Z=-2.37, p=0.018, 95% CI 0.2-1.9 lower), and more likely to make them feel they could succeed at being healthy (Z=3.44, p=0.001, 95% CI 0.6-2.1 higher) than F7. This finding is also consistent with recommendations in prior work [5,8].

The preference toward positive framing held true even when the comparison represented the same information, as in Figure 4. People found positive framing (e.g., "*you walked more than 70% of people*") more appropriate and less judgmental than negative (e.g., "*30% of people walked more than you*"). p91 noticed she had this preference "*phrasing it this way is somehow much less judgmental than when it was presented the opposite (66% of other Fitbit users did better)*", finding "*66% of Fitbit users doing better than me*

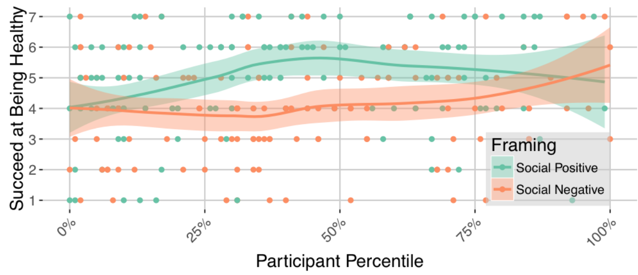


Figure 4. Participants found social comparisons framed positively made them feel more like they could succeed at being healthy than those framed negatively. The difference in preference appears to diminish at the extremes, where participants were substantially better or worse than others.

was a harsh reality." The difference in preference appears to diminish at the extremes, where participants were either better or worse than a large majority of wearers of Fitbit.

Contrast 6: Participants preferred social comparisons when they performed well relative to their peers. These participants tended to be those who have more collected data.

After normalizing to consider a participant's strict percentile (e.g., X% for F6, 100-X% for F7), higher percentiles were correlated with ratings that were more informative (Z=3.34, p<0.001, 95% CI 0.007-0.027 higher per percentage point of comparison), less judgmental (Z=-2.69, p=0.007, 95% CI 0.005-0.03 lower), and feeling more ability to succeed at being healthy (Z=2.76, p=0.007, 95% CI 0.004-0.03 higher).

In our dataset, participants with more data tended to have better comparisons. Participants with *long and consistent use* had higher comparisons than the other groups (t_{259}=4.07, p<0.001, 95% CI 4.0-15.8 higher percentage points), with no significant difference found between groups of participants with *short* and *intermittent use* (p=0.36). Future work should explore creating appropriate yet supportive comparisons for people that have a small amount of data.

Contrast 7: Participants who thought tracking was *useless or negative* preferred positively-framed social comparisons to all other framing approaches.

Figure 5 highlights this contrast. Participants who thought tracking was *useless or negative* rated positively-framed comparisons (F6) higher for feeling that they could succeed at being healthy (Z=3.16, p=0.01, 95% CI 0.2-2.4 higher on a 7-point Likert scale) than all other framings. Results were inconclusive about whether these comparisons were more appropriate (p=0.088, 95% CI -0.1-2.4) than all other framings. This suggests an opportunity to inform lapsed trackers even if they do not want to return to tracking. p34 wondered "*what percentile rank I would be in comparison to others,*" and p49 wanted to "*compare the top 5% of your step days with other users' top 5%.*"

However, participants were overall divided on whether social comparisons were appropriate. p65 described themself as "*SUPER competitive*" and found "*knowing how I compare to other Fitbit users… is really motivational!… I want that 62%*

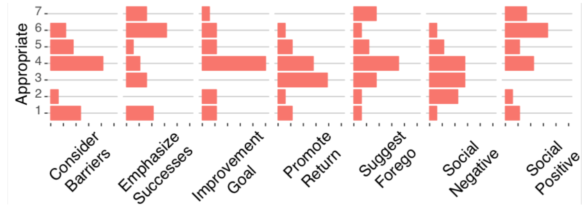


Figure 5. Participants who thought tracking was *useless or negative* still found positively-framed social comparisons appropriate.

to be 99%!" p35 saw a particularly low social comparison "*that means I'm worse than 95% of other Fitbit users?*" but found it beneficial: "*although this can be a little judgmental, it helps me anchor my perception to reality.*" However, p39 and 17 others would rather not see comparisons at all: "*I dislike being compared to other users… I'd rather see data that compared me to me.*" Participants who did not want to see social comparisons had varied perspectives on their tracking experience: 6 felt *guilt or frustration*, 8 had *conflicting feelings*, 4 thought tracking was *useless or negative*. Further research should consider how to identify self-trackers who might appreciate social comparisons.

DISCUSSION

Our findings suggest specific design considerations for personal informatics tools, such as considering a person's use pattern in selecting a visual representation of their data. Other findings suggest future research opportunities, such as rethinking representations for complex self-tracked data.

Immediate Design Recommendations

Designers can tailor personal informatics feedback, visual or otherwise, according to a person's use pattern. Participants with less data preferred more aggregated cuts through their data, while participants with more data preferred cuts that highlighted their long use. This may be because people with long use better understand their habits, and so they benefit more from designs that surface long-term trends. Short-term trackers may still be learning about their routines and thus could benefit from designs that reveal their day-to-day habits.

This paper has primarily considered the content of feedback, and our results suggest exploring these approaches in different delivery modalities. For example, many self-tracking apps email weekly reports, and our results suggest it could be effective to alter these reports for a person who has lapsed. Notification-based reminders can increase logging behavior [3], but designers might instead use notifications to provoke a person to revisit data they have collected. Our results suggest people might perceive notifications that invite them to explore their data as less aggressive and more engaging than notifications that presume a desire to resume tracking.

Recommendations Enabled by Knowing Perspective

Exploring automatic detection of tracking perspective could enable feedback tailored to provide encouragement that is more appropriate to that individual. Although we did not find strong correlations between our three participant use patterns and participant perspectives on tracking, there may be other signal to identify certain perspectives or experiences. For example, a person whose device is lost or broken is more

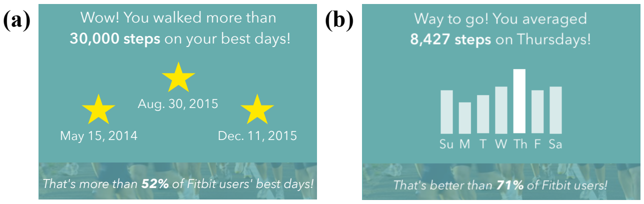


Figure 6. Designs can surface different information for lapsed trackers who do and do not want to return to tracking.

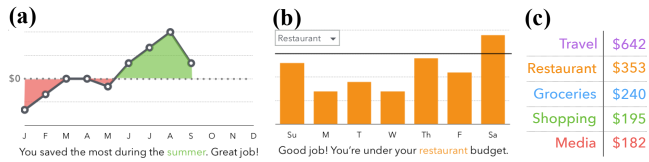


Figure 7. Lapsed financial trackers with (a) more collected data can revisit their spending habits aggregated by month, while people with less collected data may benefit from (b) seeing day-of-week trends or (c) their expenses summarized holistically.

likely to abruptly stop collecting data, while a person who loses interest may slowly decrease their collection over time. Absent such automatic detection, it may be sufficient to prompt a person to identify how they feel about their experience tracking, then vary representation of their data according to their response. One context for this might be when beginning to use a related tool or service. For example, a person beginning to use a food journaling app might be prompted regarding what other tracking they have attempted and whether they are interested in resuming step tracking.

Designers need to be cautious and consider the perspective of a person who has lapsed in tracking, or designs may cause a person to remember their tracking experience more negatively and more saliently than they would have without an additional prompt. Some people may not want to track anymore [7,11]. Others may not be able to increase their activity because of recurring health problems "*I have rheumatoid arthritis and sometimes my joints hurt too bad for me to walk*" (p57) or simply because "*I was too busy to use my Fitbit*" (p4). Designs that help people revisit and learn from past tracking without presuming a goal of returning to tracking or becoming healthier are more likely to be acceptable and create positive feelings among more people. For example, we designed Figure 6a for someone who does not want to return to tracking. It highlights the tracker's best walking days to promote reminiscence and positively compare the best days to others, leaving the tracking experience on a positive note. Care must still be taken that reminiscing may be uncomfortable for some people, depending on why they stopped (e.g., p57's arthritis above).

Participants who felt guilt or frustration over lapsing and wanted to return to Fitbit were encouraged by the visual cuts and framings presented. Figures 6a and 6b present examples of positively framed designs informed by the representations we studied, contrasting two approaches varied on use pattern. Figure 6a emphasizes high activity days drawn from a long tracking history, while Figure 6b highlights the day of the week a tracker averaged the most steps, an approach that can be effective with even a relatively short tracking history.

Extensions to Other Domains of Self-Tracking

This paper has primarily explored and discussed design for people who have lapsed in tracking physical activity using a Fitbit. Our findings can likely be adapted in other domains in which people track their behaviors, sometimes surfacing new design challenges and research questions to consider.

For example, many people self-track personal finances, with primarily manual tools (e.g., Quicken) or more automatic tools (e.g., Mint). Although many people are able to consistently track their finances for years, others abandon the practice in weeks or months [13]. Designs can adopt some of the principles we have examined with step count data. A person who tracked their finances consistently for a long time may prefer a holistic view of how their finances changed over time (e.g., Figure 7a's view of how much money they spent or saved by month). This may avoid presenting obvious habits and better motivate a person toward a savings goal. However, this representation would be barren for a person who tracked for only a few weeks. Informed by representations explored here, one alternative would be day-by-day spending relative to a discretionary budget (Figure 7b). Another might be a table divided by spending category (Figure 7c). Both of these aim to provide meaningful feedback based on a relatively short period of tracking.

Although we have focused on data that can be presented in familiar charts, some data domains offer additional challenges in aggregation and presentation. A person who tracked their location for understanding their habits or for reminiscing may find it helpful to see different cuts or framings of their location history on a map or accompanied by photos [12]. Similar aggregation and presentation questions arise in domains such as photo-based food journals and multi-dimensional lifelogs. Our results suggest additional future research informing design for people who have lapsed in tracking across diverse domains.

CONCLUSION

We contribute recommendations for designing for people who have lapsed in personal tracking, informed by contrasts identified in how participants responded to different visual cuts and framings of their data. Designs should be tailored to a person's prior experiences with tracking. A person who has tracked consistently is more likely to understand their daily and weekly habits, and may therefore prefer seeing longer-term representations of their data. Representations should also align with a person's perspective on tracking, especially to be respectful of whether a person is interested in returning to tracking. Although we explored the design of representations specifically for Fitbit data, our design recommendations inform and likely extend to other tracking in other domains.

ACKNOWLEDGMENTS

We thank Dominik Moritz, Anne Ross, Jessica Schroeder, and Kanit Wongsuphasawat for their support in developing visualizations, Matthew Kay for advice on the second analysis as well as Predrag Klasnja, Amanda Swearngin, Kyle Thayer, and Jacob O. Wobbrock for feedback on early drafts. This work was sponsored in part by the Intel Science and Technology Center for Pervasive Computing, Nokia Research, the AHRQ under award 1R21HS023654, and the NSF under awards OAI-1028195, SCH-1344613, and IIS-1553167.

REFERENCES

1. Adrienne Andrew, Gaetano Borriello, and James Fogarty. (2007). Toward a Systematic Understanding of Suggestion Tactics in Persuasive Technologies. *Proceedings of the International Conference on Persuasive Technology (Persuasive 2007)*, 259–270. http://doi.org/fj7r83
2. Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. (2013). Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *ACM Transactions on Computer-Human Interaction*, *20*(5), 1–27. http://doi.org/f23f6c
3. Frank Bentley and Konrad Tollmar. (2013). The Power of Mobile Notifications to Increase Wellbeing Logging Behavior. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2013)*, 1095−1098. http://doi.org/f2zs59
4. Michael Bostock and Jeffrey Heer. (2009). Protovis: A Graphical Toolkit for Visualization. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, *15*(6), 1121–1128. http://doi.org/b6wxzm
5. Eun Kyoung Choe, Bongshin Lee, Sean A. Munson, Wanda Pratt, and Julie A. Kientz. (2013). Persuasive Performance Feedback: The Effect of Framing on Self-Efficacy. *Proceedings of the American Medical Informatics Association Annual Symposium (AMIA 2013)*, 825–833. http://www.ncbi.nlm.nih.gov/pubmed/24551378
6. 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. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2014)*, 1143−1152. http://doi.org/bbpd
7. 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. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, 647-658. http://doi.org/bdsp
8. Sunny Consolvo, Predrag Klasnja, David W. McDonald, Daniel Avrahami, Jon Froehlich, Louis LeGrand, Ryan Libby, Keith Mosher, and James A. Landay. (2008). Flowers or a Robot Army? Encouraging Awareness & Activity with Personal, Mobile Displays. *Proceedings of the International Conference on Ubiquitous Computing (UbiComp 2008)*, 54–63. http://doi.org/dk9hqb
9. Felicia Cordeiro, Daniel A. Epstein, Edison Thomaz, Elizabeth Bales, Arvind K. Jagannathan, Gregory D. Abowd, and James Fogarty. (2015). Barriers and Negative Nudges: Exploring Challenges in Food Journaling. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2015)*, 1159–1162. http://doi.org/bbdt
10. Daniel A. Epstein, Alan Borning, and James Fogarty. (2013). Fine-Grained Sharing of Sensed Physical Activity: A Value Sensitive Approach. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2013)*, 489–498. http://doi.org/bdsq
11. Daniel A. Epstein, Monica Caraway, Chuck Johnston, An Ping, James Fogarty, and Sean A. Munson. (2016). Beyond Abandonment to Next Steps: Understanding and Designing for Life after Personal Informatics Tool Use. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2016)*, 1109-1103. http://doi.org/bjjd
12. Daniel A. Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean A. Munson. (2014). Taming Data Complexity in Lifelogs: Exploring Visual Cuts of Personal Informatics Data. *Proceedings of the ACM Conference on Designing Interactive Systems (DIS 2014)*, 667–676. http://doi.org/bbng
13. Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. (2015). A Lived Informatics Model of Personal Informatics. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, 731–742. http://doi.org/bdsr
14. Chloe Fan, Jodi Forlizzi, and Anind Dey. (2012). A Spark Of Activity: Exploring Informative Art As Visualization For Physical Activity. *Proceedings of the ACM Conference on Ubiquitous Computing (UbiComp 2012)*, 81–84. http://doi.org/bdss
15. FitBit. http://www.fitbit.com/
16. B. J. Fogg. (2002). Persuasive Technology: Using Computers to Change What we Think and Do. *Ubiquity (5)*, 89–120. http://doi.org/d8xcc5
17. Susannah Fox and Maeve Duggan. (2013). Tracking for Health. *Pew Internet*, 1–32. http://www.pewinternet.org/Reports/2013/Tracking-for-Health.aspx
18. Brendan J. Frey and Delbert Dueck. (2007). Clustering by Passing Messages Between Data Points. *Science*, *315(5814)*, 972–976. http://doi.org/fcq25h
19. Thomas Fritz, Elaine M. Huang, Gail C. Murphy, and Thomas Zimmermann. (2014). Persuasive Technology in the Real World: A Study of Long-Term Use of Activity

Sensing Devices for Fitness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2014)*, 487–496. http://doi.org/bbdv

20. Jon Froehlich, Leah Findlater, and James Landay. (2010). The Design of Eco-Feedback Technology. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010)*, 1999–2008. http://doi.org/cbsjjr
21. Rúben Gouveia, Evangelos Karapanos, and Marc Hassenzahl. (2015). How Do We Engage With Activity Trackers? A Longitudinal Study of Habito. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, 1305–1316. http://doi.org/bdst
22. Daniel Harrison, Paul Marshall, Nadia Bianchi-Berthouze, and John Bird. (2015). Activity Tracking: Barriers, Workarounds and Customisation. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, 617–621. http://doi.org/bdsv
23. Geert H. Hofstede. (1997). *Cultures and Organizations: Software of the Mind*. McGraw-Hill, New York.
24. Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. (2015). Personal Visualization and Personal Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, *21*(3), 420–433. http://doi.org/bdsw
25. International Data Corporation. (2016). The Worldwide Wearables Market Leaps 126.9% in the Fourth Quarter and 171.6% in 2015, According to IDC. http://www.idc.com/getdoc.jsp?containerId=prUS41037416
26. Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. (2011). How to Evaluate Technologies for Health Behavior Change in HCI Research. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)*, 3063–3072. http://doi.org/dzgdm6
27. Amanda Lazar, Joshua Tanenbaum, Christian Koehler, and David H. Nguyen. (2015). Why We Use and Abandon Smart Devices. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, 635–646. http://doi.org/bdsx
28. Min Kyung Lee, Junsung Kim, Jodi Forlizzi, and Sara Kiesler. (2015). Personalization Revisited: A Reflective Approach Helps People Better Personalize Health Services and Motivates Them To Increase Physical Activity. *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015)*, 743–754. http://doi.org/bkrs
29. Ian Li, Anind Dey, and Jodi Forlizzi. (2010). A Stage-Based Model of Personal Informatics Systems. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2010)*, 557–566. http://doi.org/bh8zsb
30. James J. Lin, Lena Mamykina, Silvia Lindtner, Gregory Delajoux, and Henry B. Strub. (2006). Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. *Proceedings of the International Conference on Ubiquitous Computing (UbiComp 2006)*, 261–278. http://doi.org/crcvd9
31. Susan Michie, Charles Abraham, Craig Whittington, John McAteer, and Sunjai Gupta. (2009). Effective Techniques in Healthy Eating and Physical Activity Interventions: A Meta-Regression. *Health Psychology*, *28*(6), 690–701. http://doi.org/bj595b
32. Susan Michie, Stefanie Ashford, Falko F. Sniehotta, Stephan U. Dombrowski, Alex Bishop, and David P. French. (2011). A Refined Taxonomy of Behaviour Change Techniques to Help People Change Their Physical Activity and Healthy Eating Behaviours: the CALO-RE taxonomy. *Psychology & health*, *26*(11), 1479–98. http://doi.org/fdg8jr
33. John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers. (2014). Personal Tracking as Lived Informatics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2014)*, 1163−1172. http://doi.org/bbdz
34. Donald A. Schön. (1983). *The Reflective Practitioner*. Basic Books. http://doi.org/fv2g6n
35. Patrick C. Shih, Kyungsik Han, Erika Shehan Poole, Mary Beth Rosson, and John M. Carroll. (2015). Use and Adoption Challenges of Wearable Activity Trackers. *iConference 2015 Proceedings*, 1–12. http://hdl.handle.net/2142/73649
36. Tableau. http://www.tableau.com/
37. Tammy Toscos, Anne Faber, Kay Connelly, and Adity Mutsuddi Upoma. (2008). Encouraging Physical Activity in Teens: Can Technology Help Reduce Barriers to Physical Activity in Adolescent Girls? *Proceedings of the International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth 2008)*, 218−221. http://doi.org/cvmffx
38. Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, and Jeffrey Heer. (2015). Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, *22*(1), 649−658. http://doi.org/bdsz