On the Integration of Self-tracking Data amongst Quantified Self Members

Mark Whooley
The University of Melbourne
Parkville Victoria 3010, Australia
mail@markwhooley.com

Bernd Ploderer
The University of Melbourne
Parkville Victoria 3010, Australia
ploderer@unimelb.edu.au

Kathleen Gray
The University of Melbourne
Parkville Victoria 3010, Australia
kgray@unimelb.edu.au

Self-tracking, the process of recording one's own behaviours, thoughts and feelings, is a popular approach to enhance one's self-knowledge. While dedicated self-tracking apps and devices support data collection, previous research highlights that the integration of data constitutes a barrier for users. In this study we investigated how members of the Quantified Self movement—early adopters of self-tracking tools—overcome these barriers. We conducted a qualitative analysis of 51 videos of Quantified Self presentations to explore intentions for collecting data, methods for integrating and representing data, and how intentions and methods shaped reflection. The findings highlight two different intentions—striving for self-improvement and curiosity in personal data—which shaped how these users integrated data, i.e. the effort required. Furthermore, we identified three methods for representing data—binary, structured and abstract—which influenced reflection. Binary representations supported reflection-in-action, whereas structured and abstract representations supported iterative processes of data collection, integration and reflection. For people tracking out of curiosity, this iterative engagement with personal data often became an end in itself, rather than a means to achieve a goal. We discuss how these findings contribute to our current understanding of self-tracking amongst Quantified Self members and beyond, and we conclude with directions for future work to support self-trackers with their aspirations.

Quantified Self, self-tracking, self-monitoring, personal informatics, data integration

1. INTRODUCTION

There is a growing interest in HCI in self-tracking technology. Self-tracking in general refers to the process of recording one's own behaviours, thoughts and feelings, which can enhance self-knowledge and foster reflection (Li et al. 2010). Many people simply track information like their weight or their exercise in their heads (Fox and Duggan 2013). However, technology offers various benefits, like automating data collection or making manual tracking less tedious. Hence, HCI researchers have explored the benefits of self-tracking technology to encourage reflection on a variety of behaviours, like electricity consumption (Pierce and Paulos 2012), transportation habits (Froehlich et al. 2009), what we eat (Smith et al. 2007), and exercise habits (Consolvo et al. 2006).

In recent years, self-tracking technologies have also become successful commercial products like the Nike+ Fuelband or the family of Fitbit products, which allow anyone to keep track of their exercise. Numerous web and mobile apps developed track everything from moods, to diets, to sleep patterns, and it has been estimated that by 2017 there will be 1.4 billion mobile sensing health and fitness app

downloads worldwide (ON World 2013). Societal trends like increased privileging of good health and the importance of taking personal responsibility for one's health (Lupton 2013) have further helped to promote self-tracking technology.

While technology has made the collection of personal data more convenient, recent research indicates that the volume and the variety of available data has also introduced new challenges (Choe et al. 2014, Li et al. 2010). There are various technical challenges, like data from many systems in multiple formats and barriers to exporting and sharing it. On the other hand, the volume and the complexity of the data also requires a set of skills to decide what data to collect, how to integrate data and how to make sense of and learn from the data.

In light of these challenges, we have turned to early adopters of self-tracking in the Quantified Self community to examine how they cope with the challenges of integrating various data. Quantified Self is an active, international community who share their knowledge and experiences both online as well as in meetings in local groups. As of February 2014, there are meetings held in 106 cities in 36 countries,

including local groups in London, Manchester, Leeds and Edinburgh (Quantified Self 2013).

In what follows, we present the findings from our study on integration amongst Quantified Self members. We start by providing background on self-tracking, Quantified Self and theory from the related field of personal informatics and we explain our methods. The findings highlight the importance of intentions behind self-tracking. They offer a detailed description of the methods used to integrate and represent data, and the opportunities that these representations offer for reflection.

2. RELATED WORK

There are many different terms in the literature that describe technologies that allow users to track personal data. These are as a result of the differing aspects the authors are wishing to emphasise, personal informatics systems (Khovanskaya et al. 2013, Li et al. 2010, Li et al. 2011) seek to highlight the overall system, whereas self-tracking (Ploderer et al. 2012b) and self-monitoring technologies (Choe et al. 2014), take their terms from historical methods that now utilise technology, whereas personal analytics (Choe et al. 2014) shifts the focus from technology to humans. The Quantified Self group is founded on the principle of 'self-knowledge through numbers' helping to popularise these technologies and actions in public discourse.

Li et al. (2010) propose a stage-based model to describe how users of self-tracking technologies obtain and transform personal data into information and self-knowledge. The first stage—preparation—occurs when people make a decision to start self-tracking, determining what and how they will collect information. Collection, the second stage is when the user collects data; this can be achieved in a short or long timeframe, depending on the extent that technology is used (Li et al. 2012). Integration is the next stage where the user prepares, combines, and transforms data into (typically visual) representations allowing analysis and reflection. The fourth stage is the reflection stage where users engage with their data to reflect on their behaviours. Action, the final stage is where users take the self-knowledge they have obtained from their self-tracking activities and apply it to their behaviour (Li et al. 2010). This process is iterative, as users adjust their self-tracking practices due to insights from the data, changes in goals, availability of data and tracking tools, etc.

There has been research on various aspects of the stage-based model. Li et al. (2012) created two personal informatics systems that required users to consider the contextual data they should gather during the preparation stage. Khovanskaya et al. (2013) explored new, primarily automatic and designer-led, methods for users during the

preparation stage. New technical methods identified facilitate the collection of alternative data sources (Khovanskaya et al. 2013, Li et al. 2012, Rawassizadeh et al. 2012). Furthermore the creation of designer-led visualisations (Ng et al. 2011) has helped our understanding of how people reflect on their data. Contextual data has been introduced to representations of data to stimulate users' reflection on their behaviours (Li et al. 2012), increasing our understanding of the reflection stage. This has led to a systems-led approach, as most of this work appears to focus on the automatic creation of visualisations. The difficulties faced by ambivalent users of self-tracking data have been explored (Ploderer et al. 2012), which provides further insight into the barriers at the action stage.

Despite the wealth of research on various aspects of the stage-based model, little attention has been paid to the integration stage, possibly due to the expectation that this is an automatic process and therefore less important. Kaipainen and colleagues (Kaipainen et al. 2011) concentrated their efforts on improving automatic integration of multiple sources of data within personal informatics systems. However, most systems developed seek to collect individual pieces of data and users are required to integrate other sources if they require them (Li et al. 2010, Li et al. 2012). While previous work points out barriers at the integration stage, i.e. having to organise data from multiple inputs and scattered representations of data (Li et al. 2010), little is known about the practices of users who have overcome these barriers and effectively integrate data for further reflection and action.

Hence, the aim of this paper was to develop a nuanced understanding of how self-tracking data is integrated into representations. Given the emphasis of previous work on automating integration (Kaipainen et al. 2011), we explored the practices of self-trackers at different degrees of automation as well as the different types of data representations that resulted from this work. Furthermore, the stage-based model emphasises a holistic approach to self-tracking (Li et al. 2010); hence we were interested in the relationship between integration and other stages in the process of self-tracking. Therefore, we explored the intentions that led Quantified Self members to track personal information and how these intentions may have shaped data integration. Additionally, we were interested in how the data representations developed during the integration stage shaped people's ability to reflect and act on the data.

3. APPROACH

In order to get a better understanding of the practices at the integration stage we analysed 51

videos from the Quantified Self website where members presented their self-tracking projects.

3.1. Data Set

The data set consisted of 51 videos posted on the Quantified Self website <www.quantifiedself.com>. The Quantified Self has 96 affiliated groups across the world (Quantified Self 2013), including four in the UK, each of which host regular 'show and tell' meetings. Presentations follow a standard format where they introduce what was done, how, and what they learned from it. At the time of this study in October 2013, the website hosted over 200 videos of these presentations. The first author reviewed the 100 most recent videos that were posted between October 2011 and October 2013 for identifying activities related to integration, such as accounts of organising data, combining data from different sources and creating visualisations and other representations. The 51 videos meeting the criterion were saved offline and selectively transcribed to capture discussions of integration.

Quantified Self members may not be a representative sample of everyday users of personal informatics systems, limiting the scope of application. The study's transferability (Miles and Huberman 1994) outside of the Quantified Self is limited as these early adopters may have different behaviours and attitudes during the integration stage. The data may also not be representative of the Quantified Self, as one would expect that members giving presentations that get posted online are the ones who have gone the extra mile. The use of videos in the analysis where users provided a narrative style presentation may have added further limitation to the findings. The presenters may have excluded aspects of the integration stage due to lack of time, potentially limiting the study's authenticity (Miles and Huberman 1994). Finally, unlike Quantified Self studies in health areas (Lupton 2013) we sought to characterise practices of self-trackers rather than evaluate the quality of their actions.

3.2. Data Analysis

The analysis was primarily qualitative in order to characterise how Quantified Self members integrated data, how their intentions shaped integration, and how their representations in turn supported reflection. Descriptive statistics were used to provide context to the qualitative findings, however the sample size was too small and the self-selection of Quantified Self members who presented and shared their videos was too specific to establish quantitative results that are applicable across a wider population. Instead, our aim was to show what is possible given the determination and dedication of Quantified Self members.

The analysis was primarily data-driven, though previous work on self-tracking, i.e. the stage-based model (Li et al. 2010), sensitised our analysis. The first author viewed all videos three times to immerse himself into the data and to capture preliminary observations that went beyond what has been reported in the stage-based model. These observations included different approaches, i.e. intentions in self-tracking, types of data representations, as well as assertions about how intentions shape data integration. The observations were abstracted and captured in mind-maps. Together with related data, these observations and assertions were written down in analytic memos (Birks et al. 2008). The observations were then reviewed, challenged, and refined in weekly meetings with the other authors, who reviewed the evidence and guided the analysis. The first author subsequently coded the transcripts to look for further evidence for the preliminary observations. This coding process was done initially using Microsoft Excel noting the timestamps that were relevant to codes and transcribed material. Once we processed key codes and identified them in the analysis we applied and counted their occurrence to provide an indication of common observations. These codes were then refined using pattern coding techniques (Miles and Huberman 1994) where relationships between codes were identified and codes were grouped together into themes. We wrote these themes up using methods advocated by Neuman (2011) including a definition of the code, a flag that helps identify them in the videos, a qualification that helps exclude aspects not relevant and sample data. Not all of the initial memos were included in the resulting findings; some were discarded due to a lack of empirical evidence in the analysis or when they went beyond the research scope.

The following section provides an overview of those themes that were prominent in the data and hence selected during the final round of coding; they describe the intentions for self-tracking, practices associated with different degrees of system-support during integration, representations produced during integration, and how these representations supported reflection. All data in the findings section have been anonymised through pseudonyms and random identifiers for videos.

4. FINDINGS

4.1. Intentions for Collection and Integration

The Quantified Self members collected a variety of personal information that could broadly be categorised as health and lifestyle data. Health data often focussed on diet and bodyweight, but also sleep quality, stress levels, diabetes (glucose levels) and fitness levels. Lifestyle tracking typically

involved tracking of daily activities, like hobbies, online behaviours and how they spent their time.

In our analysis we found that these diverse self-tracking activities were driven by two different intentions: self-improvement and curiosity. Out of the 51 presenters in our analysis, 27 could be characterised as being driven by self-improvement. These Quantified Self members presented themselves as being driven by a specific goal, and they used technologies to quantify and analyse both the factors that contributed to their goal as well as the outcomes themselves. For example, Gérard's goal was to reduce his weight to a certain level. He used different technologies to collect data about fitness activities and his diet, as well as scales to keep track of his weight.

"I am a huge fan of technologies that makes tracking fitness activities as efficient as possible, if it's easy and it's fun chances are I am going to try it....using RunKeeper every time I go for a workout and with Weight Watchers you know tracking what I am eating throughout the day" (Video #2595, 8:58).

An alternative intention for self-tracking that was identified during the analysis was curiosity. This occurred when people presented themselves as being intrigued about their self-tracking data. These 24 self-trackers were not guided by a desire to improve. Instead, they had open-ended questions that they sought to explore through self-tracking. While these Quantified Self members typically presented self-tracking as a means to answer open-ended questions, our observations indicated that they were often opportunistic and examined data that were readily available, simply out of curiosity rather than a well-articulated question. In some instances it appeared that analysing personal data was almost an end in itself rather than a means to answer a pre-defined question. This curiosity was well illustrated in video #8825 where Peter discussed how he developed an app that helped him find out his modes of transportation over a period of one month. This example illustrates how users can be driven to explore themselves without necessarily intending to strive for self-improvement.

"Around the middle of August I started asking myself how do I get around? I don't own a car, I walk a whole lot, I take the bus, I do some other stuff...I just one day got this idea in my head, I want to get a better sense of it, it is something that is really routine" (Video #8825, 0:24).

4.2. Degrees of Automation in Data Integration

The stage-based model (Li et al. 2010) suggests that integration is system-driven, or involves user-driven methods or a combination of both. Based on these categories, the following sections characterise how

Quantified Self members integrate data with different degrees of system support.

4.2.1. System-driven Integration

Entirely system-driven integration methods require no involvement by the user during the integration stage and a preconfigured representation is created for the user, ready for the reflection stage.

"All these data streams are flowing into the [device] where they can be synchronised and data can be spit out in the form of web pages as a result of analysis of those different data streams" (Video #3470, 2:00).

System-driven integration is designer-led and provides a clear objective as to what is required in the representation. This integration method would be considered quick as it seeks to harness the power of technologies, requiring little effort or time on the part of the user, allowing them to concentrate on reflection activities. Of the 13 videos analysed that used automatic methods, only one indicated curiosity intention, illustrating that automatic integration often occurs when the user is outcome orientated as these users know what they are looking for in the data. By conducting automatic integration the user effectively skips any input into this phase and jumps directly from collection to the reflection stage where a representation of their data has been automatically created.

4.2.2. User-driven Integration

User-driven integration on the other hand requires a level of involvement by the user during the integration stage. Users take time to analyse the data collected and ascertain the best way to create a representation that will assist them during the reflection phase. In video #6409 Mohammed took sleep data and cognitive test scores in an effort to ascertain a correlation. This was significant effort on Mohammed's part and took a considerable amount of time. The example also shows how Mohammed tried to combine multiple system-driven integrations into a master set of data.

"I wanted to combine my Zeo data; to see if that could predict my quantified mind data. To do that I created this thing called the aggregate mental performance, which is a simple average of all the different quantified mind tests...Working with data can be hard, trying to get the different formats from the different devices like the Zeo versus the quantified mind and getting the different programs to work together can be a hassle...And also it is important to do a sanity check, I got lucky and only lost two hours instead of two or three days because I did a simple graph and realised I aligned my data incorrectly at first" (Video #6409, 4:02).

User-driven integration typically occurred when users wanted to combine multiple sources of data they had collected, which generally required

significant effort and time. They devised methods that would allow them to identify if mistakes were being made by cross referencing data and validating it as they merged and analysed it. Sometimes the analysis portion was straightforward; with basic statistical and analytical skills users could take multiple representations and compile them into a dashboard view. However, other user-driven integrations were more complicated as users who had the skills wrote software programs that could extract data and integrate it into a representation they were happy with. It became apparent in our analysis that users, who were driven to track reflexively out of curiosity, were more interested in integrating the data manually. Of the 23 videos indicating curiosity intentions 22 employed user-driven integrations.

4.2.3. Integrating Complementary Data

In some of the videos the presenters commented on how they integrated other pieces of information that they did not explicitly track. This involved a moment where they would recollect something that helped explain some of the data that were illustrated in the representation created. In the example below Alfred was tracking his weight and noted spikes in his weight gain that coincided with interruptions to his usual daily routine.

"There are a couple of plateaus...my habit is...key. Anything that interrupts my routine will cause a notable effect on my weight...my birthday was last weekend and my mother made cake and how can you say no to your mother when she presents you with vegan chocolate iced double-layered cake? So that caused one heck of a spike" (Video #3626, 8:17).

This could be seen as a point of feedback for the user to validate the representation they were creating. It can include an informal addition of contextual data that the user is aware of, but has not made explicit on the representation created. This data can often be used as an explanation of anomalies identified in the data.

4.3. Representations

The videos analysed illustrated three orientations that representations can take: binary, structured and abstract representations. These were evident after users sought to manually integrate their data in a manner that would allow them to extract sufficient knowledge from what they had collected.

4.3.1 Binary Representations

Binary representations distil the data integrated into one of two results, for example whether enough exercise has been achieved over a period of time or not, providing an immediate representation to the user. They can take many formats that provide instant feedback, for example a sound or flashing

light. Determined to exercise more, Hari devised a method that would allow him to integrate real-time data to create a visual prompt allowing him to immediately correct his behaviour. He created a cycle seat (figure 1) that enabled him to work at his computer and cycle at the same time. A red light under the computer would turn on if he stopped peddling for a sustained period of time; he has built in an algorithm that does allow some breaks. This binary representation is a simple distillation of the data a user collected, generally in real-time, to one of two options (in Hari's case a red light on or off). Although we only found three instances of binary representations, these representations are primarily for users motivated by self-improvement.

Figure 1: Cycle seat with binary representations where a red light appears if the user does not cycle enough within a time period (Video #7499, 3:18)

Binary representations could be seen as very simple forms of structured representations. We discuss binary representations separately, because unlike other structured representations they were typically set up to support very brief reflection and instant changes in action.

4.3.2. Structured Representations

Structured representations are well-known representations such as tables and graphs, which show some form of relationship between two or more variables. Examples of simpler structured representations include bar graphs, pie charts, or scattergrams; and more complex ones included data matrixes, statistics, or maps. Structured representations were the most common form of representation in our analysis, occurring in 41 of 51 videos. These representations were used both by people who appeared to be motivated by a specific outcome and those motivated by curiosity.

People used structured representations to tell a story of their data over a time period. In video #7980 Andrew created a scattergram of data he gathered from a t-shirt he wore that had sensors to track his posture (figure 2). He was able to use this data to tell how his posture changed during a typical day, identifying times when his posture was better than other times.

Figure 2: Scattergram, an example of a simple structured representation that required the user to plot their data (Video #7980, 5:03)

More complex structured representations required further work on behalf of the user to get their data in a format that will give them the knowledge they require. In video #8324 Samantha analysed data about her diet and weight using statistical methods, producing a t-score of her data, to verify what she had found within a scattergram representation previously created. Such complex representations generally required the user to work with their data more than just once to develop a good understanding of it.

4.3.3. Abstract Representations

Six Quantified Self members in our sample created representations that lacked an obvious structure. We characterised these representations as abstract, because similar to abstract art, the relationship between the representations of these Quantified Self members and their references to the source data was not easily recognisable for the audience. In fact, these Quantified Self members saw self-tracking as an opportunity for artistic expression. They created representations that were ambiguous because they invited engagement and allowed for multiple interpretations by the audience. For example, Tomasz wanted to create artistic representations of his surfing and skateboarding experiences. Using motion sensor data he converted numerical data collected on X, Y and Z-axes into abstract representations as illustrated in figure 3 below. The use of abstract representations can be seen as self-interpretations of the data and involve an element of creativity from the user during the integration stage. These representations become digital artefacts of his experiences.

Figure 3: Abstract representations of self-tracked activity do not immediately point to what the data are and they are open to interpretation or require input from the user to explain their meaning (Video #4224, 1:31)

"My notion is this: is there a way I can forge a relationship with the statistical data, information that I have gathered from various types of interactions and use it as a form of expression... One of the things I am very interested in is surfing and skateboarding...I developed a system that allows me to take data from those activities and create things like [figure 3]" (Video #4224, 0:45).

4.4. Influencing Reflection

The intention for collecting data and the representations produced during the integration stage shaped consecutive reflection in various ways. Not surprisingly, Quantified Self members with clear intentions for self-improvement typically created binary and structured representations to reflect on the various factors influencing the goal they wanted to achieve. For example, Gérard (video #2595), whom we spoke about before, was interested in reducing his weight. To assist with this goal he tracked his calories (both food intake and exercise) and his weight. He then aligned his overall calorie intake with his weight loss. The representations he chose provided clear visibility of the metrics he was tracking, allowing for immediate concentration on these aspects.

Quantified Self members who created abstract representations were typically driven by curiosity. Their abstract representations allowed them to explore various characteristics of their data that may not have been otherwise apparent to them. For example Andrew (video #1737) tracked the books he read using images of each book cover over time. Through that process he found that, although the number of books he read had increased, the literary quality of the books had decreased. This outcome was only apparent to him when he saw the images of the books.

Furthermore, the videos illustrated that complex structured representations as well as abstract representations evolved iteratively. Initial insights from the reflection stage led people to collect additional tracking data as well as to re-work their initial representations. For example, Samantha (video #8324) discussed how she created a visual representation of the factors contributing to her weight and fitness, paused and found that she was unable to derive meaning from her initial attempt. Hence she rearranged the data and ran further statistical tests to develop a better understanding of how she had arrived at her current fitness level.

"The first thing everyone does, we make some pretty charts. We make some data visualisation, so I did that and I like the first chart. It's looking pretty good, it's going down...But the second chart, ok, so they kind of look the same. I don't know if there is really that big of a difference... So I don't know what to make of this...I'm going to do some t-tests" (Video #8324, 3:44).

Figure 4: Users were found to create a dashboard of representations that helped provide a full picture of their self-tracking activities (Video #4157, 7:14)

Some users, identified in 18 videos, produced multiple representations of the same set of data which allowed them to have a dashboard of information available for reflection. This allowed them to look at their self-tracking data from multiple angles providing a more holistic picture. Users with programming skills often added an element of interactivity to their dashboard, allowing them to drill into information of interest. For example, Tanya created a website of multiple representations detailing every movie she went to over a period of eight years (figure 4). She created numerous structured representations of the data including graphs displaying how many movies she saw a year, a scattergram of her movie attendance over time against a third party movie rating, and how much movies cost. She combined all this information in a dashboard and used her programming skills to satisfy her curiosity by allowing herself to drill into specific aspects, for example movie titles or dates.

5. DISCUSSION

The findings have provided a detailed analysis of the ways in which Quantified Self members integrate personal data. While previous work highlighted various challenges that self-trackers face at the integration stage, i.e. the integration of multiple data sources (Li et al. 2012), and made attempts to reduce them by automating integration with the support of the self-tracking system (Kaipainen et al. 2011), we have shown how early adopters and self-tracking enthusiasts manage these challenges with their personal data.

In particular, this study has produced three findings. Firstly, the findings have highlighted two different intentions—striving for self-improvement and curiosity in personal data—which shaped the ways in which Quantified Self members integrated data. Individuals who tracked personal data out of curiosity often integrated their data manually because they wanted to explore what novel insights that data could offer them. Individuals striving for self-improvement, on the other hand, sought systems that automated data integration. For them integration was a means to achieve their goals, for example to become fitter. Secondly, this study has

provided a detailed account of the activities involved in integrating personal data; we have shown how users integrated data with different degrees of system support and we have described three methods for representing data—binary, structured and abstract representations. Abstract representations were typically the result of explorations by individuals with a curiosity in personal data, whereas people with either intention used binary and structured representations. Thirdly, the findings have shown how the intention behind self-tracking and the representations that the Quantified Self members produced has shaped consecutive reflection. Automation allowed individuals to skip integration and continue with reflection on their actions, whereas manual integration was typically an iterative process, where the self-trackers moved back and forth between creating representations and reflecting on their action. In the following sections we will discuss how these findings extend our current understanding of data integration in the context of self-tracking.

5.1. Intentions for Self-tracking and Integration

The findings showed that Quantified Self members were self-tracking either for self-improvement or simply out of a curiosity in personal data. Similar to previous studies of Quantified Self members (Choe et al. 2014, Li et al. 2011), our analysis showed that self-tracking is often driven by a desire to improve one's health or some other lifestyle aspect, e.g. how efficiently one spends their time.

Li et al. (2011) also report that many self-trackers are simply exploring data to specify their goals and/or to identify the factors that may influence their lives. These intentions did not surface in our study, possibly because our analysis was limited to videos of show and tell presentations posted on the Quantified Self website, which are possibly skewed towards successful projects with a clear narrative. For similar reasons, we did not encounter self-trackers who highlighted the collection of rewards or their interest in gadgets as their primary intention as reported in related work (Rooksby et al. 2014). Interviews would be better suited to find members with these intentions.

However, we found that a significant subset of the Quantified Self members (24 of the 51 videos) in this study collected data out of a curiosity in what the data may reveal. Unlike the users in Li et al.'s study (2011), these Quantified Self members were not exploring the data to establish goals or improve their lives. Instead they self-tracked to get a better understanding about a facet of their everyday lives that they otherwise would be unaware of. Some members like Peter, who tracked his modes of transportation, were simply interested in better understanding their own activities, without wanting to change them. Rooksby et al. (2014) framed this

style as documentary tracking. For members like Tomasz, on the other hand, self-tracking appeared to be an end in itself. Like a craftsman (Sennett 2008), Tomasz regarded the tracking of his skateboarding as a form of engagement to do a job well for its own sake. Unlike traditional craftsmen, however, the material he worked with was primarily digital in the form of personal data.

The two intentions of self-improvement and curiosity in turn shaped how the Quantified Self members integrated their personal data. Members striving for self-improvement sought systems that automatically integrated data from various sources for them. They manually integrated their data when the technology did not adequately support them with their goals and when they were sufficiently motivated and skilled. However, as pointed out by Li et al. (2011), these members benefit from systems that automatically integrate data for them and allow them to effectively jump from the collection to the reflection stage, because their goal is to learn about themselves to improve their lives.

Quantified Self members who tracked personal data out of curiosity, on the other hand, often integrated their data manually because they wanted to explore what novel insights that data could offer them. Driven by this curiosity, they combined several sources of data, which required manual work. Some members devised their own software to have more control over how to track, integrate and visualise their data. While this may seem like a difficult and work-like practice, members like Tomasz (who visualised his skateboarding) did not regard it as such. As reported in HCI studies on craftsmanship (Ploderer et al. 2012a, Rosner and Ryokai 2009), these members had an aspiration for quality and they enjoyed their engagement with their personal data for its own sake. Unlike other self-trackers, people driven by curiosity do not necessarily seek technology that automates integration, because it can diminish opportunities for engagement with personal data. However, these self-trackers benefit from tools that allow them to export their data so that they can interrogate it freely without the constraints of any given tool.

5.2. Different Approaches to Integration

In addition to the processes of automatic and manual data integration discussed above, the findings have also highlighted how Quantified Self members integrate qualitative data to account for events and trends in their numerical data. This is important, as previous work pointed out that a common mistake amongst self-trackers is that self-trackers focus too much on tracking symptoms without capturing contextual information to interpret these symptoms (Choe et al. 2014). Furthermore, this is important; Lupton (2013) warns of the allure of self-quantification, which deceptively reduces the

complexity of health and lifestyle factors into simple, seemingly scientifically neutral numbers. The findings showed that while metrics were important to help self-trackers see how they made progress, they used qualitative data to assist them in developing a more holistic understanding of their health and/or lifestyle.

Furthermore, we described three methods for representing data during the integration process: binary, structured and abstract representations. Structured representations like tables and line charts illustrate relationships between two or more variables. These representations are well covered by previous work, both in terms of how to produce representations that support precise, effective and quick analysis (Tufte 1983) as well as the popularity of these representations in Quantified Self practice (Choe et al. 2014).

Binary and abstract representations, on the other hand, have not yet surfaced in the discourse around the Quantified Self. Binary presentations were a simple form of structured representation, where data was distilled into one of two results, typically to provide real-time feedback to users whether they were on track with their goals or not. As illustrated in the cycle seat, binary representations required work to set up, but afterwards they automatically integrated and fed back data to the users. Binary representations were similar to persuasive technology that allows users to remain aware of their goals and progress, e.g. in the context of saving energy (Pierce and Paulos 2012) and encouraging exercise (Consolvo et al. 2006).

Abstract representations were depictions of self-tracking data where the relationship between the data and its representation were not obvious to the viewer. As illustrated in the findings, these representations took inspiration from the visual arts. They were forms of self-expression, as illustrated in Tomasz' visualisation of skateboarding. Furthermore, these representations were made deliberately ambiguous. Abstract representations may seem counterintuitive in the context of Quantified Self, because as illustrated so far, a key aim in quantifying health and lifestyle information is to provide unambiguous feedback for reflection and self-improvement. However, as is well known in HCI (Gaver et al. 2003), the ambiguity of representations can also be seen as a strength because it can encourage close personal engagement with and reflection on the data.

5.3. Data Representations and Reflection

The last contribution of this work is to extend our current understanding of different forms of representation and the opportunities they offer for subsequent reflection. As discussed above, abstract representations were deliberately ambiguous, which encouraged close personal engagement with the

data. Furthermore, these abstract representations were typically refined over time, where users moved back and forth between integrating and visualising data and reflection on the data. This immersion in the data through iterative integration and reflection is another facet of the craftsman-like engagement with personal data which highlights a desire to do the job well for its own sake (Sennett 2008).

Reflection on structured representations is well documented in related work. Dourish and Mazmanian (2011) discuss the way that different representational forms provide different structures for reflection and self-knowledge. For example, lists support ordering of information whereas tables show relations between different data points. Furthermore, Li et al. (2012) point out that the amount of data has a bearing on people's ability to reflect on their data. In particular, automated tracking and integration enables people to track more information, which can make reflection in the short-term more difficult but provides benefits in the long-term to develop a better understanding of the factors that influence one's goals. We made similar observations in the videos where users found it difficult to sift through large quantities of data in the short-term. Typically, the Quantified Self members in this study created multiple representations through their dashboard to deal with this challenge, and they refined them over time as their understanding of their personal data developed.

Finally, binary representations could be seen as supporting 'reflection-in-action' (Schön 1983). In his work about reflective practice, Schön differentiated between reflection-in-action and reflection-on-action, where the first denotes contemplation at the time of doing ("thinking on our feet"), whereas in the latter case reflection takes place after actions are completed. Structured and abstract representations typically supported reflection-on-action, where data collection, integration and reflection were conducted one after another. Binary representations, on the other hand, provided feedback for reflection at the time of action. For example, the cycle seat presented in the findings provided feedback at the user's desk to support thinking about and to encourage moderate exercise. This representation supported the user's knowing-in-action about the importance of exercising to counterbalance the health effects of sedentary work while he was going about his work.

6. CONCLUSION

This study has provided detailed insight into how Quantified Self members integrate a multitude of data, a process that is often overlooked as it is assumed that the self-tracking system will automate this task for the user. We found data integration a rich area of activity amongst Quantified Self

members, often one that is engaging and critical to their overall experience with self-tracking. In particular, we have described how people integrate self-tracking data, what representations they produce and how these representations influence reflection. We found that the intention behind self-tracking shaped how people integrated their data. People who had clear goals to improve an aspect of their lives indeed sought to minimise the effort involved in integrating data, whereas people who displayed a curiosity in their personal data found integration a critical part of their overall engagement with their personal data. Automating this part of the self-tracking process would diminish their experience. However, one area for future work is to design and evaluate self-tracking tools that allow people to export and combine their data freely to better support their curiosity in personal data. Given the large amount of data that not only Quantified Self members but most of us collect, whether it be through email communication or our engagement with social media, tools that allow us to explore our digital footprint likely would be of interest to the general public.

Furthermore, we described in detail how Quantified Self members integrated their data, the representations they produced, and how these representations shaped subsequent reflection. Most members followed a linear process where they collected, integrated and reflected upon their personal data, often in an iterative manner. Some members used self-tracking as self-expression, using abstract, ambiguous representations to encourage close engagement with the data. A small group of Quantified Self members devised representations that supported reflection-in-action. These representations provided them with immediate feedback on the actions that interested them. We see particularly the area of supporting reflection-in-action as another promising avenue for future work, which has the potential to expand the benefits of self-tracking. Rather than asking people to embark on long self-tracking projects to reflect on their lifestyles, well-designed personal devices and services have the potential to offer immediate feedback on activities that we find interesting and thereby enrich our everyday life without adding more work to it.

7. REFERENCES

- Birks, M., Chapman, Y. and Francis, K. (2008) Memoing in qualitative research: Probing data and processes. *Journal of Research in Nursing*, 13(1), pp. 68-75.
- Choe, E. K., Lee, N. B., Lee, B., Pratt, W. and Kientz, J. A. (2014) Understanding Quantified-Selfers' Practices in Collecting and Exploring Personal Data. in *CHI 2014*: ACM. pp. 1143-1152.

- Consolvo, S., Everitt, K., Smith, I. and Landay, J. A. (2006) Design requirements for technologies that encourage physical activity. in *CHI 2006*, New York: ACM. pp. 457-466.
- Dourish, P. and Mazmanian, M. (2011) Media as Material: Information Representations as Material Foundations for Organizational Practice. in *Process Symposium 2011*, Corfu, Greece. pp. 1-24.
- Fox, S. and Duggan, M. (2013) *Tracking for Health* Washington, DC: Pew Internet and American Life Project.
- Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B. and Landay, J. A. (2009) UbiGreen: Investigating a mobile tool for tracking and supporting green transportation habits. in *CHI 2009*, New York: ACM. pp. 1043-1052.
- Gaver, W. W., Beaver, J. and Benford, S. (2003) Ambiguity as a resource for design. in *SIGCHI 2003*, Ft. Lauderdale, Florida, USA: ACM. pp. 233-240.
- Kaipainen, K., Honka, A. and Saranummi, N. (2011) Personalized behavior change support for disease prevention. in *EMBS 2011*, United States: IEEE. pp. 880-883.
- Khovanskaya, V., Baumer, E. P. S., Cosley, D., Volda, S. and Gay, G. K. (2013) "Everybody Knows What You're Doing": A Critical Design Approach to Personal Informatics. in *CHI 2013*, Paris, France: ACM. pp. 3403-3412.
- Li, I., Dey, A. and Forlizzi, J. (2010) A stage-based model of personal informatics systems. in *CHI 2010*: ACM. pp. 557-566.
- Li, I., Dey, A. and Forlizzi, J. (2011) Understanding my data, myself: supporting self-reflection with ubicomp technologies. in *UbiComp 2011*: ACM. pp. 405-414.
- Li, I., Dey, A. and Forlizzi, J. (2012) Using context to reveal factors that affect physical activity. *Transactions on Computer-Human Interaction*, 19(1), pp. 1-21.
- Lupton, D. (2013) Quantifying the body: monitoring and measuring health in the age of mHealth technologies. *Critical Public Health*, pp. 1-11.
- Miles, M. B. and Huberman, A. M. (1994) *Qualitative data analysis: an expanded sourcebook*, Thousand Oaks: Sage Publications.
- Neuman, W. L. (2011) *Social research methods: qualitative and quantitative approaches*, 7th ed., Boston: Ally & Bacon.
- Ng, T., Ou Jie, Z. and Cosley, D. (2011) pieTime: Visualizing Communication Patterns. in *PASSAT 2011 / SocialCom 2011*: IEEE. pp. 720-723.
- ON World (2013) *Mobile Sensing Health & Wellness*, Available: <http://onworld.com/mobilesensing/health/> [Accessed 18 November 2014].
- Pierce, J. and Paulos, E. (2012) Beyond energy monitors: interaction, energy, and emerging energy systems. in *CHI 2012*, Austin, Texas, USA: ACM. pp. 665-674.
- Ploderer, B., Leong, T., Ashkanasy, S. and Howard, S. (2012a) A process of engagement: engaging with the process. in *DIS 2012*, Newcastle Upon Tyne, United Kingdom: ACM. pp. 224-233.
- Ploderer, B., Smith, W., Howard, S., Pearce, J. and Borland, R. (2012b) Things you don't want to know about yourself: ambivalence about tracking and sharing personal information for behaviour change. in *OzCHI 2012*, Melbourne, Australia: ACM. pp. 489-492.
- Quantified Self (2013) *Quantified Self*, Available: <http://quantifiedself.com/> [Accessed 18 August].
- Rawassizadeh, R., Tomitsch, M., Wac, K. and Tjoa, A. M. (2012) UbiqLog: a generic mobile phone-based life-log framework. *Personal and Ubiquitous Computing*, 17(4), pp. 621.
- Rooksby, J., Rost, M., Morrison, A. and Chalmers, M. C. (2014) Personal Tracking as Lived Informatics. in *CHI 2014*. pp. 1163-1172
- Rosner, D. K. and Ryokai, K. (2009) Reflections on craft: Probing the creative process of everyday knitters. in *C&C 2009*, New York: ACM. pp. 195-204.
- Schön, D. A. (1983) *The reflective practitioner: how professionals think in action*, New York: Basic Books.
- Sennett, R. (2008) *The craftsman*, London: Yale University Press.
- Smith, B. K., Frost, J., Albayrak, M. and Sudhakar, R. (2007) Integrating glucometers and digital photography as experience capture tools to enhance patient understanding and communication of diabetes self-management practices. *Personal and Ubiquitous Computing*, 11(4), pp. 273-286.
- Tufte, E. R. (1983) *The Visual Display of Quantitative Information*, Cheshire, Connecticut: Graphics Press.