Behaviour & Information Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tbit20

Designing a personal informatics system for users without experience in self-tracking: a case study

Amon Rapp, Alessandro Marcengo, Luca Buriano, Giancarlo Ruffo, Mirko Lai & Federica Cena

BEHAVIOUR & INFORMATION TECHNOLOGY, 2018
VOL. 37, NO. 4, 335–366
https://doi.org/10.1080/0144929X.2018.1436592

Designing a personal informatics system for users without experience in self-tracking: a case study

Amon Rapp[a], Alessandro Marcengo[b], Luca Buriano[b], Giancarlo Ruffo[a], Mirko Lai[a] and Federica Cena[a]

[a]Department of Computer Science, University of Torino, Torino, Italy; [b]TIM, Torino, Italy

ABSTRACT
Thanks to the advancements in ubiquitous and wearable technologies, Personal Informatics (PI) systems can now reach a larger audience of users. However, it is not still clear whether this kind of tool can fit the needs of their daily lives. Our research aims at identifying specific barriers that may prevent the widespread adoption of PI and finding solutions to overcome them. We requested users without competence in self-tracking to use different PI instruments during their daily practices, identifying five user requirements by which to design novel PI tools. On such requirements, we developed a new system that can stimulate the use of these technologies, by enhancing the perceived benefits of collecting personal data. Then, we explored how naïve and experienced users differently explore their personal data in our system through a user trial. Results showed that the system was successful at helping individuals manage and interpret their own data, validated the usefulness of the requirements found and inspired three further design opportunities that could orient the design of future PI systems.

ARTICLE HISTORY
Received 22 May 2017
Accepted 29 January 2018

KEYWORDS
Human–computer interaction; user studies; visualisation systems and tools; information visualisation; Personal Informatics; Quantified Self

1. Introduction

The opportunities offered by wearable and ubiquitous technologies open new horizons for Personal Informatics (PI) systems, defined as 'those that help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge' (Li, Dey, and Forlizzi 2010, 558). A plethora of personal data can be currently collected automatically, from physiological parameters and psychological states to performance values and habits.

This act of collecting data on one's own behaviour is known as self-monitoring in clinical psychology (Elliot et al. 1996). Conceived originally as an assessment method to gather information about behaviours undetectable by an external observer (Foster et al. 1999; Korotitisch and Nelson-Gray 1999), it has been employed as an intervention technique because of its reactivity, i.e. the fact that the process of behaviour recording produces a change in behaviour itself (Miltenberger 2007). PI systems improve the act of self-monitoring by allowing individuals to track data everywhere at any time, potentially enabling the arising of self-reflection and triggering processes of behaviour change. At the end of the last century, PI systems started to be used by researchers and technology enthusiasts, such as the members of the Quantified Self community (Marcengo and Rapp 2014).

We recently witnessed the spreading of a variety of ad hoc wearable devices and applications aimed at monitoring users' behaviours in the consumer market. Adding to this, advancements in wearable and ubiquitous technologies could make the gathering of personal data less obtrusive in the near future, by automatically detecting a variety of states that now still rely on the conscious act of self-reporting. Such advancements could foster the adoption of PI tools among a larger audience of potential users, making their personal data readily available to them (Rapp et al. 2015, 2017). These 'potential users' might be curious regarding PI to discover something about themselves: however, they might not have impellent needs to self-track, as well as knowledge about using trackers effectively. Kim, Paulos, and Mankoff (2013) identified two stages of how people adopt technology: in the first stage, interest, engagement and fun are the main motivations, whereas in a later stage the instrument is integrated into daily routines (Huang et al. 2015). Reflecting on how we could design better trackers to engage individuals who show interest in self-monitoring, but do not have any prior experience in this practice, becomes then essential to favour the adoption of PI tools among a wider user base.

In this new landscape, a question arises: should PI systems be rethought in order to better fit the needs and desires of individuals who are curious about PI tools and who have not tried them yet? Despite the favourable

CONTACT Amon Rapp amon.rapp@gmail.com

market forecasts (IdTechEx 2014), there is an increasing scepticism about the real capabilities of PI devices in providing users with real benefits (Hunter 2014; Maddox 2014), especially those who are not expert in dealing with 'numbers' (Hilviu and Rapp 2015; Rapp et al. 2016). A recent research stressed that one-third of the Americans who purchased a tracker abandoned it after only six months (Ledger and McCaffrey 2014). Despite the success of PI devices among specific categories of individuals (such as the Quantified Selfers, who are keen on tracking per se [Choe et al. 2014], and users motivated to change their behaviour [Fritz et al. 2014]), such a phenomenon suggests that people find it difficult to engage with tracking technologies.

Moving from this issue, our research aims at finding how PI systems can be better integrated into everyday life, by identifying specific barriers that may prevent their widespread adoption and finding new solutions to overcome them. The main contributions of this article are the following: (i) it proposes a new PI system that embeds five specific user requirements aimed at engaging and satisfying users without prior experience in self-tracking; (ii) it validates such requirements through a qualitative user study, showing how naïve and expert users differently value the system; and (iii) it proposes three further design opportunities that could orient the design of future PI tools. The paper is structured as follows. Section 2 presents related work, while Section 3 describes the user requirements we found through the initial study. Section 4 outlines the system we developed according to such requirements and Section 5 presents the evaluation study. Finally, Section 6 highlights the limitations of our work, while Section 7 closes the paper with conclusions and future work.

2. Background and related work

Pioneering examples of PI applications can be found in life logging research projects (Cheng, Golubchik, and Kay 2004; Mann 2004; Gemmell, Bell, and Lueder 2006). As years passed, other academic research appeared, aimed at designing systems to collect and visualise personal information for therapeutic and rehabilitation purposes, or for promoting behaviour change (e.g. Consolvo et al. 2006, 2008a, 2008b; Matthews and Doherty 2011; Kay et al. 2012). As of today, the availability on the market of a plethora of self-tracking applications and devices has raised the acceptance of PI (Rapp and Cena 2015). Smartphone applications – such as Moves (movement) and Expereal (mood) – or ad hoc wearable devices – like the Jawbone UP bracelet (sleep, physical activity, mood and food) – track users' behaviours, automatically or through self-reporting, and feed information back to them. Furthermore, research in wearable and ubiquitous technologies prefigures a future when a variety of data, which now need to be self-reported or require obtrusive technologies, could be automatically detected by non-invasive devices (Cena, Likavec, and Rapp 2015). For example, Kawamoto, Hiroyuki, and Seiki (2013) use a wrist-worn accelerometer to detect respiratory rate and then estimate REM sleep, the subjective level of drunkenness, fever and smoking cessation (Kawamoto, Tanaka, and Kuriyama 2014). Lu et al. (2012) track stress states from human voice using smartphones. Sumida, Mizumoto, and Yasumoto (2013) estimate physical load and its variation during walking with the already available functions of a smartphone. FoodBoard (Pham et al. 2013), an instrumented chopping board, uses optical fibres and embedded camera imaging to identify and track ingredients during food preparation on its surface.

2.1. PI systems

Despite these technological advancements, it is still unclear whether PI systems could provide users with sufficient benefits to be integrated into their everyday life. Could these systems offer a meaningful reflection of the user's self? Will she be able to manage, explore and interpret the data provided by these technologies in a simple way?

It has been highlighted that the data that most PI instruments collect are siloed (Rapp and Cena 2014). This limits their applicability, narrowing the vision on the user's 'self', by focusing on separated aspects of her life (Rapp and Tirassa 2017). In the last few years, different commercial systems have been developed to import data from different sensors/applications into a single user account, attempting to provide a wider vision on the personal collected data (see Table 1). Nevertheless, these platforms do not mash-up data, offering just a devoted space for each sensor/application, with no opportunity of directly comparing data or easily discerning patterns over time (Bentley et al. 2013). Moreover, information is displayed through numbers and stats: a variety of research showed that many individuals lack the ability to understand data from graphs (Galesic and

Table 1. Personal data aggregators.

	Google Fit	Apple Health	Tictrac	Fluxtream	Heads Up Health	Beeminder
Main purpose	Fitness data	Fitness/health data	Health data	Personal data	Fitness/health data	Goals/productivity data

Garcia-Retamero 2011) and statistical data (Ancker and Kaufman 2007).

To overcome these issues, several prototypes have been designed in the academic field in order to collect and merge personal data, providing insights on how information can be made more meaningful. Ubifit (Consolvo et al. 2008a), for example, targets individuals with the need to increase their regular physical activity and consists of a glanceable display, an interactive application and a fitness device: the display uses the metaphor of a garden that blooms throughout the week as physical activities are performed. Although the system focused only on physical activity, it showed how we may enhance self-reflection by designing intuitive metaphors.

Salud! (Medynskiy and Mynatt 2010), instead, aggregates data from different sources leaving users free to interpret them by using time series graphs and tables. However, it could be difficult for users that are not able to manage quantitative information and abstract visualisations to gain insights from their data. Although Salud! made a step forward in the direction of a greater integration among different data sources, it raises the question about how we can make all these interconnected data more understandable by users.

Li, Dey, and Forlizzi (2012) designed a PI system prompting contextual information in order to support reflection on the factors that affect physical activity. This system allows users to recollect the context related to a specific event when trying to display physical data. We were inspired by the possibility of using contextual information to elicit memories connected with specific data sets: this work, in fact, suggests to further investigate ways for providing users with a richer experience.

Finally, Bentley et al. (2013) designed a health mashup system that highlights connections between diverse wellbeing data through natural language, stressing how this could support self-awareness that may lead to behaviour change. This work puts into question the traditional way of displaying quantitative information, encouraging us to explore new modalities to provide users with meaningful insights when using PI tools.

2.2. Personal visualisation's works

Personal Visualisation, as the design of interactive data representations for use in a personal context (Huang et al. 2015), and casual Information Visualisation, namely those visualisations addressing the broad range of user populations with a special focus on casual audience (Pousman, Stasko, and Mateas 2007), have been recently receiving increasing attention. Their focus on personal data and visualisations for everyday situations might provide insights for PI design.

Feltron's Annual Reports (2014), a series of annual pamphlets made by Nicholas Feltron, artistically transform personal data into a variety of diverse info graphics and tables, correlating different aspects of his daily life, such as food eaten, songs listened and places visited. This work shows how personal data can be displayed by using impressionistic representations aimed at giving a snapshot of one's life, rather than by employing core infovis techniques, such as treemaps and scatterplots.

LastHistory is an interactive visualisation for displaying personal music from last.fm (Baur et al. 2010). It allows users to analyse data, with an emphasis on temporal patterns, and to remember a life event in the past, by pairing listening histories with contextual information, such as personal photos and calendar entries. Similar to Li, Dey, and Forlizzi's (2012) work, it stresses the importance of context for recollecting events tied to certain data.

ShutEye (Bauer et al. 2012) provides a peripheral display on the wallpaper of the user's mobile phone to support awareness about suggested activities that might promote good sleep quality, prompting key information as needed. It also uses simple bar charts for suggesting a timeline for specific activities, without requiring any explicit interaction from the user.

Epstein et al. (2014), instead, designed different visualisation modalities to represent cuts, i.e. subsets of collected PI data with some shared feature. They found that the visual cuts that people value vary dramatically, recommending future designs to offer a variety of them. However, they did not find a correlation between participant goals and the cuts they valued most or least.

Finally, Huang, Tory, and Bartram (2016) integrated quantitative personal data into a personal digital calendar. They aimed to make the data readily accessible and more comprehensible, by aligning Fitbit data and life events. By leveraging users' calendars, the visualisation provided context for reasoning about temporal fitness data.

As we have seen, market and academic research is oriented to find both new opportunities for integrating different data sources into a unique platform and new visualisations for interacting with personal information. However, it is still not clear whether these systems could also satisfy the needs of users who are curious about PI tools but who have not tried them yet. As we have seen in the Introduction, the increasing sales of self-tracking devices may hide a fad, since recent research highlighted that new users quickly abandon their trackers (Hunter 2014; Ledger and McCaffrey 2014; Maddox 2014). This may be due to the fact that such instruments are not designed for individuals who do not have experience in self-tracking. Actually, there is a lack of research on this kind of users and we do not even know precisely what needs they may have when collecting and interacting with personal data.

2.3. PI user studies

Most descriptive studies of how people use commercially available tools were focused on experienced trackers and the problems they face. Li, Dey, and Forlizzi (2010) first explored real-world practices with PI tools, highlighting how experienced users find different barriers when managing personal information: these users find it difficult to integrate different data sources and reflect upon them; while other barriers are represented by issues in retrieving, exploring and understanding information, e.g. due to the fact that data are scattered in different locations. Based on their findings, Li, Dey, and Forlizzi (2010) define a PI model that is composed of five sequential stages through which people transition when dealing with personal data: *preparation*, *collection*, *integration*, *reflection* and *action*. Li, Dey, and Forlizzi (2011) further studied the needs of users that use self-tracking instruments to change their own behaviour, emphasising how they are interested in visualising data that may reveal their current status with regard to a behaviour change goal, in figuring out what goals would be appropriate to pursuit, or in determining how different factors may affect their behaviour. These users are motivated to actively manage the collected data, by employing, for example, paper graphs or by reviewing data logs.

This proactive attitude was also found among the Quantified Selfers (Choe et al. 2014). This extremely experienced user group tracks multiple data simultaneously with the intention of identifying correlations and causation among them. They are mainly focused on numbers, preferring line and bar charts to visualise what they have collected. Although they are not exempt from encountering barriers when using PI tools, they find solutions by building their own tracking tools or by creating customised visualisations, showing that a strong motivation in understanding data might overcome fails and shortcomings of the currently available instruments. Fritz et al. (2014) confirmed a focus on behaviour change in people engaged in the long-term usage of activity trackers, highlighting how they are extremely attached to their devices.

As we may see, the practices and needs of experienced trackers have been deeply investigated by the HCI community, particularly focusing on behaviour change and technological matters. However, minor attention has been given to the practices and needs of inexperienced users. More specifically, Rooksby et al. (2014) claimed the need to shift the attention from technical problems to be resolved to a range of lived activities where people use personal information and find its meaning. In doing so, they included in their study sample participants who never (or barely) used an activity tracker before. This study is important because it emphasises that most people tie tracking activities to their lives, hopes, worries, careers and so on. Users may track for different reasons, like *collecting rewards*, documenting activities (*documentary tracking*), pursuing a behaviour change goal (*directive tracking*), finding correlations (*diagnostic tracking*) or pure interest in technology (*fetishised tracking*). Although these findings moved the attention to a wider variety of users and motivations to track, Rooksby et al. involved only four novices in their research, and their discourse remained focused on the description of participants who had already used at least one tracker before the start of the research.

Lazar et al. (2015), instead, interviewed users who are not willing or able to integrate 'smart devices' in their lives. Interestingly, their participants abandoned almost 80% of the devices purchased for the study. Although participants were initially motivated by the novelty of the device and the curiosity about it, for many of them a drop-off effect occurred as novelty diminished and the cost of maintaining the device became irritating. This study is important because it suggests that naïve users may have specific needs and desires. Building on these findings, we think that it is crucial to further explore the particular needs that they may have, with a greater focus on PI matters. In fact, Lazar et al. were more interested in exploring how 'smart devices' (i.e. wearables) were used, without connecting them with the PI discourse. For this, they did not consider PI apps, not relying on a physical dedicated device. Moreover, they risked to bias the sample by selecting only people with a focus on technology.

To summarise, Table 2 provides a snapshot of the main works cited in this section, pointing to their goals, strengths and weaknesses. As a matter of fact, it is important to investigate inexperienced users for different reasons: (i) they represent a new category of potential users that could largely expand the market of PI tools; (ii) they are understudied in the current literature; (iii) recent research highlights that they quickly abandon PI devices; and (iv) research also emphasises that this could be due to a lack of engagement, since data are found not useful, and their maintenance is too cumbersome (Lazar et al. 2015). Therefore, finding novel design requirements specific for them, and building systems addressing such requirements, might deal with the quick disengagement that they experience when interacting with PI devices.

Table 2. Related work summary.

PI systems System	Goals	Strengths	Weaknesses
Consolvo et al. (2008a)	To encourage users in doing more physical activity.	Glanceable display and metaphoric visualisation to support users in the ongoing activity.	Focus on a single kind of data (i.e. steps). Single visualisation modality.
Medynskiy and Mynatt (2010)	To support users in self-managing their health condition.	Aggregation of data from different sources. Openness to third-party development.	Focus on abstract visualisations (graphs and numbers).
Li, Dey, and Forlizzi (2012)	To encourage users in doing more physical activity.	Usage of contextual information to enrich the interpretation of data.	Focus on a single kind of data (i.e. fitness data). Focus on textual information.
Bentley et al. (2013)	To support users in changing behaviour.	Aggregation of data from different sources. Emphasis on correlations among different data. Use of natural language.	No flexibility in connecting new sources. Focus on health data. No graphic representations of data.
Personal visualisations **Visualisation**	**Goals**	**Strengths**	**Weaknesses**
Feltron (2014)	To give a snapshot of Feltron's life	Different views on the same data. Impressionistic visualisations of data.	No interactivity.
Baur et al. (2010)	To explore and remember listening histories	Temporal and value queries to explore historical data. Locating listening events is supported by synchronised photos and calendar entries.	Exclusive focus on listening data. No spatial information is given to recollect a listening event.
Bauer et al. (2012)	To recommend activities that promote good sleep quality	Peripheral display to support users in the ongoing activity.	Exclusive focus on sleep data.
Epstein et al. (2014)	To help self-trackers identify meaningful and actionable findings	Diverse 'data cuts' that leverage different representations.	No insights for connecting specific visualisations to specific goals are provided.
Huang, Tory, and Bartram (2016).	To make personal visualisation applications fit into everyday life routines	Intuitive view to relate data to calendar activities.	No contextual information about location is given.
PI Studies **Study**	**Goals**	**Strengths**	**Weaknesses**
Li, Dey, and Forlizzi (2010)	To explore real-world self-tracking practices.	They highlighted barriers in self-tracking. They defined a model of PI use.	No novices in self-tracking involved in the study.
Li, Dey, and Forlizzi (2011)	To understand how users reflect on their data.	They described how self-trackers manage their data and what kind of questions they ask about their data.	No novices involved in the study.
Choe et al. (2014)	To understand Quantified Selfers' practices.	They described how an 'extreme user group' successfully uses self-tracking devices.	Results are not applicable to the wider population due to the Quantified Selfers' idiosyncrasies.
Fritz et al. (2014)	To explore long-term usage of PI devices among fitness trackers	They studied how users use trackers in 'in the wild' contexts.	No novices involved in the study.
Rooksby et al. (2014)	To explore tracker users' motivations to self-monitor	They highlighted different motivations to track apart from behaviour change. They discovered that tracking activities are tied to users' lives.	Only four novices involved in the study. Focus on the description of experienced users.
Lazar et al. (2015)	To explore practices of disengagement among PI users	They described why novices abandon their devices. They suggested that novices may have specific needs and desires.	They did not frame their results in the PI discourse. They exclusively involved employers of a technology company. They did not study PI apps.

3. Diary study

To investigate how inexperienced users perceive and use PI tools we conducted a diary study. This is a longitudinal method of understanding users' behaviour in their everyday life: they differ from other field study techniques because participants have control on the aspects of the behaviour recorded and the timing of recording (Carter and Mankoff 2005). Diary studies allow the researcher to collect data in contexts that may be hard to observe 'because of social or physical reasons' (Church and Smyth 2009). They have been employed to investigate the use of a number of different technologies, such as SMSs (Grinter and Eldridge 2003), tools to support multitasking (Czerwinski, Horvitz, and Wilhite 2004) and mobile devices (Sohn et al. 2008). While an extensive presentation of the study results can be found in Rapp and Cena (2016), in the following we will outline how this research yielded a series of specific user requirements.

3.1. Procedure

We recruited 14 participants (age $M = 31.9$, $SD = 10.1$, females = 8) through recruiting emails and snowball sampling. We screened participants using a telephonic interview, by investigating their interest and curiosity towards PI (see Appendix A.1 for the screening criteria).

We then selected participants with no previous experience in PI, but who showed a sense of curiosity and interest towards the possibility of discovering something about themselves and using PI tools. We ascertained this aspect by asking users whether they ever thought to buy or install a tracker, whether there were aspects of their behaviour that they wanted to understand better (such as sleep, levels of physical activity) and whether they thought that this goal could be reached by using a self-tracking tool. None of them had any pressing need for gathering data about their psychological or physical condition, such as pathological conditions (mood disorders, chronic diseases, etc.) or a strong desire to change their behaviour. We considered this kind of participants interesting for our purposes because they could represent a new potential user base for PI: even without having any established a need or goal to collect these data, they could find their own ends during the exploration and use of personal information, becoming interested in knowing something about themselves.

We defined this type of users as *naïve users* (i.e. *inexperienced users*) in contrast to the 'experienced users' investigated in the previously cited studies. Another way to look at this class of users is to frame them in the Transtheoretical Model of Health Behaviour Change (Prochaska and Velicer 1997), which inspired Li's model of PI, by stating that these individuals could conceivably be in a precontemplation or contemplation stage with regard to the utility of self-tracking: our participants, in fact, did not think about it (4), or they had an ambivalent attitude towards it by being interested in it, but at the same time they did not yet decide to adopt this practice (10). All participants owned a smartphone and were comfortable with technology, but most of them (10) were not working in an ICT company or studying technology disciplines: they were undergraduate students (4), PhD students (2), post-doctoral researchers (3),[1] psychologist (1), commercial operators (2),[2] designer (1) and lawyer (1).

We divided users into two groups: seven of them (age $M = 32.1$, SD = 12.7, females = 4) had to use a Jawbone Up bracelet, recording data for sleep, food, physical activity and mood (P1–P7). The other seven participants (age $M = 31.7$, SD = 7.8, females = 4) had to use three PI applications on their smartphones, such as Moves, Dreamboard, Sleepbot, MyFitnessPal, T2 Mood Tracker, Daytum, Expereal, recording data for movements, dreams, sleep, food, daily tasks and mood (P8–P14). We differentiated participants in these two groups because we wanted: (i) to allow participants to collect data on different aspects of their life; (ii) to investigate both the usage of PI wearable devices and PI applications; and (iii) to explore differences between managing different data scattered among different tools (the apps), and grouped in a unique interface (Jawbone Up).

As to the first group, we selected the Jawbone Up, after a desk analysis, since it is popular, not bulky and capable of tracking and integrating a variety of information (sleep, physical activity, mood and food intakes). Other similar devices, such as FuelBand, are exclusively focused on physical activity, while others, such as FitBit, are of larger dimensions, therefore more obtrusive. For the second group, apps were chosen after a heuristic evaluation to exclude usability problems. An initial set of applications was selected to grant variety in terms of types of parameters/behaviours tracked. Three independent experts, who all had previous experience in heuristic evaluations, assessed the apps by using heuristics for mobile applications (Bertini, Gabrielli, and Kimani 2006). After the analysis, the experts compared their lists and reconciled analysis. When it was not possible to unanimously determine whether a specific issue was severe or not, the decision was taken by the majority. Two apps (MoodPanda and FitDay) were excluded because the experts identified major usability problems in the apps' key features, following Nielsen's (1994) severity ranking scale. Instead, the flaws identified in the selected apps were all judged as cosmetic problems or minor usability problems. Exceptions are represented by Daytum and Sleepbot, where experts also identified major usability problems, which, nevertheless, were not considered crucial for the correct daily working of the tools, as they involved only ancillary features.

Participants were allocated to the two groups depending on their curiosity in the aspects to be monitored and their predisposition towards trying a wearable device (e.g. two participants thought that the Jawbone Up would not have fit with their personal dressing style, and they were assigned to the second group). Applications were assigned by taking into account the participants' interest in a specific state to be monitored, and the overall balance between automatically recorded data and manually inserted data.

Participants assigned to the first group (P1–P7) were given a Jawbone Up and asked to install its app on their smartphone. Moreover, they were required to wear the bracelet night and day. Participants allocated to the second group (P8–P14) were asked to install the apps keeping them active for the entire period of the research. All participants were invited to self-report the data required by the assigned tools and to display their daily data regularly. Participants had to use the assigned tools for a minimum of 10 days. After this period, they could choose whether to end their participation at any point or to continue up to one month, as we wanted to investigate their spontaneous engagement in the tools provided. One participant continued the trial for 1 month. The average period of involvement was 17 days. Users were not compensated for their participation.

The researchers provided each user with some primary written instructions in an informative leaflet, and an electronic diary in a Microsoft Word™ file: the document was made of a series of tables, one for each day of the research, where rows represented the different elements we wanted to be collected (see Appendix A.1 for the diary prompts). Participants were asked to fill in the diary every day both in the evening, by recollecting their reflections and emotions related to the tools, and right after an important episode, by taking notes of the momentary thoughts triggered by the situation. Participants showed to have a high compliance with the assigned tasks. Failures in filling the diary were distributed in the same way throughout the study period (with no peaks in its last phases). Therefore, the decreasing involvement that some participants showed in the course of the research (e.g. forgetting to insert information in the tools) was probably connected with a fading interest towards the provided instruments, rather than the study itself (they kept filling in the diary, even when they did not use the tools). We interviewed participants on the day after their decision to quit the study. The interviews were semi-structured and lasted one hour (see Appendix A.1 for the list of the questions). We audio recorded and then transcribed the answers, analysing them along with the data coming from the diaries through a thematic analysis (Braun and Clarke 2006).

The analysis was inductively oriented. Findings were coded independently by the first author and the last author who generated initial codes: data were broken down by taking apart sentences and paragraphs and by labelling them with a name. The results were then compared for consistency of text segmentation and code application. A detailed segment-by-segment review, including assessments of consistency in defining the beginning and end of segments and the application of codes within segments (MacQueen et al. 2008), was carried out. All inconsistencies were discussed and resolved. Such codes were then grouped into categories independently by the two researchers and labelled. The themes identified were finally compared and reviewed to solve inconsistencies.

A complete presentation of the study and its findings goes beyond the scope of this paper. An extensive description of the study's results along with a series of design strategies for PI can be found in Rapp and Cena (2016). In that work, future-oriented design recommendations were outlined, mostly emphasising lines of research that still need to be explored. Here, instead, we need to focus on how the study's findings could be immediately used to ground the design of a PI system. Among the recommendations defined in there, we thus selected and reworked those that we think essential to increase the value of the collected data (requirements 2 and 3). We further defined three novel user requirements, not previously described, capable of solving current relevant issues in PI (requirements 1, 4 and 5) and implementable in the short term. As an example of how coding led to such requirements, MOTIVATION, STYLE, AESTHETIC, ENGAGEMENT, EXPLORATION, ANALYSIS, INTERACTION, FUN have been grouped into the category 'Visualisation modalities' which yielded the definition of Requirement 4 'Offer different views on data'; whereas SUPPORT, CONNECTION, CAUSATION, VARIATION, SYNTHESIS, HETEROGENEITY, INTERPRETATION were grouped into the category 'Relations among data' that turned into Requirement 5 'Highlight data correlations'.

We will now proceed to outline the defined requirements, informing them with both a summary of our study's results and previous literature.

3.2. User requirements

3.2.1. Requirement 1: Promote the integration of different sources of data

Participants (13 out of 14) were bewildered by the dispersion of data among different applications. Users who were using the mobile apps, like P11 and P12, found it difficult to figure out how the different kinds of collected data could be integrated into a meaningful frame. They would have wanted a unique place to find all their information, not only to increase the meaningfulness of the gathered data, but also for privacy and security reasons. This problem was not artificially created by providing participants with incompatible tools. In fact, this need of integration goes beyond the interconnection of different instruments, pointing to a singular and secure place to group all the personal data collected over time, apart from their origin. Moreover, this need was also stressed by the participants assigned to the first group. Although they were partially satisfied by the possibility of visualising the tracked information in a single interface (i.e. that of Jawbone app), they emphasised how the kinds of data coming from the tool were not flexible. P3 and P4, for example, highlighted how the instrument was unable to add other data coming from new sources.

This requirement suggests providing users with the general picture of the information they gathered, allowing for the interconnection of data coming from different channels. PI tools should leave the users free to choose and add new data sources according to their needs, by designing flexible platforms that can homogenise different formats and simplify data interoperability. Integration barriers were noticed in previous studies as well. Li, Dey, and Forlizzi (2010) emphasised how experienced users encountered problems when data came from multiple inputs, or when reflection happened in multiple outputs. Liu, Ploderer, and Hoang (2015) further noted that sleep-trackers experience difficulties in finding proper tools to integrate or export their sleep data to combine them with information coming from other sources. The importance of integration of data was also stressed by Bentley et al. (2013). Choe et al. (2014) and Whooley, Ploderer, and Gray (2014), instead, showed how extremely experienced users, such as the Quantified Selfers, try to overcome these issues by manually integrating data. Different from the users involved in these studies, our participants showed a more 'passive' attitude: they did not actively search for a solution capable of overcoming the encountered problems, asking instead for a more proactive role of the systems in integrating diverse data sources. Moreover, this requirement points to the importance of having a unique 'actor' in charge of guarding users' data, and of designing for flexibility when creating data aggregators.

3.2.2. Requirement 2: Support users in remembering their data

Our participants signalled a lack of contextualisation of the information provided. Ten participants (out of 14), like P2, were unable to recollect the events connected to a particular information: for example, the reasons behind a peak in her physical activity three days before. This resulted in a diminished value of the collected information. Although four participants (out of 14), like P3, thought that PI tools could be useful for remembering the past, the lack of contextual cues prevented them from recollecting their memories. Six participants (out of 14), like P4 and P8, suggested that anchoring data to the spatial dimension could improve the act of remembering. Moreover, Moves users reported that they were almost always able to recollect the memories tied to the amount of steps covered in a specific moment in the past.

This requirement suggests supporting users' episodic memory (Tulving 2002), by providing them with a map for re-experiencing their data. Together 'when' and 'where' can favour the reminiscence of past events linked to the information gathered, which, in turn, can make the collected data more valuable and meaningful.

The importance of context when visualising PI data was also noted by Li, Dey, and Forlizzi (2012). They showed how contextual information can positively impact on wellbeing and behaviour change. Epstein et al. (2014) further suggested that showing visual cuts of PI data appropriate for the user's current context may increase opportunities for behaviour change. Here, instead, we want to focus on spatial contextual cues, apart from their effect on behaviour: a spatial map, in fact, can be employed to increase the perceived value of the gathered data making users remember what happened when they were collecting a particular information.

3.2.3. Requirement 3: Support users in identifying with their data

Our participants noted how the assigned tools provided too abstract visualisations: although 4 participants (out of 14) appreciated the employment of stats and graphs to display quantitative information, most of them perceived their own data to be far removed from what they considered their 'self'. As a result, participants were unable to develop an emotional bond with their data, showing a scarce engagement with the assigned tools. P3, for example, reported how she felt 'distant' from her data, which were 'cold' and 'anonymous', preventing her from being emotionally involved in their exploration. Furthermore, 6 participants (out of 14) wished for more concrete visualisations, capable of supporting identification with the collected data. P9, for example, emphasised how the identification with her data would have helped her in making them more valuable.

This requirement recommends presenting the user's data in an intuitive form, fostering her identification with them. Impressionistic and concrete images of data in which a user can recognise herself may enhance her engagement in using PI systems, increasing their capability of heightening self-awareness and improving the perceived importance of the collected data. They, in fact, may have positive outcomes in terms of enjoyment, immersion and positive affect (Birk et al. 2016), helping users feel their data to be closer to themselves.

The need to engage users in data visualisations was also observed by Bentley et al. (2013): they argued for the employment of natural language to make the data more meaningful, as numbers and graphs could be hard to understand for a large part of the population. Gouveia, Karapanos, and Hassenzahl (2015) further emphasised the importance of user engagement to sustain behaviour change, by turning glances at data into engage sessions, through a well-crafted and tailored story. While these works aimed at using linguistic forms to involve users in their own data, in this requirement we point to the importance of designing intuitive and concrete graphical representations. In fact, they can support identification and emotional connection with data, not only increasing the user engagement, but also making her feel closer to her own information.

3.2.4. Requirement 4: Offer different views on data

Nine participants (out of 14) showed decreasing engagement in visualising and exploring their data as the study proceeded. They mostly displayed some information, but stopped doing so at the first difficulties. Participants argued that a unique way of presenting the data may not be suitable to engage all types of users. P3, P9 and P11, for example, would have appreciated both a holistic view, 'immediate and direct', and a detailed view, 'for those who have time for studying their data'. P5, P8 and P13 stopped visualising their data in the last days of the study, reporting that they were annoyed by the fact that the tools 'showed always the same things', as said by P13.

This requirement highlights the diverse needs that drive users in exploring their own data. Providing different visualisation modalities, which can satisfy both the users who want to have an immediate and quick access to their data and the ones who would like to go deep in the understanding of what they gathered, is a way to favour the spreading of PI tools to a larger audience. Epstein et al. (2014) recommended offering a variety of 'visual cuts' to address individuals' idiosyncrasies. Our results further emphasise the need for designing at least two different visualisation modalities, one more direct and intuitive, the other one more detailed and comprehensive.

3.2.5. Requirement 5: Highlight data correlations

All participants wanted to find correlations between different kinds of data. However, they often encountered difficulties in interpreting the displayed information: discovering useful insights required long time and strong efforts. Moreover, 11 participants (out of 14) showed not having sufficient motivation to deeply explore their data. Instead, they would have wanted greater support when seeking the most important relations among the collected information: P1, who was not capable of detecting relevant correlations among the gathered data, reported the need for help from the Jawbone Up itself. Others bumped into the same problem: even though they intended to find how their information could intertwine, they were unable to do so through the features provided by the tools. As a result, the perceived utility of these instruments decreased during the study. Our participants did not have previously established objectives in mind on how to use this technology, but they were curious and interested in it. However, this curiosity rapidly faded away, as the tools were unable to prompt them with easy-gaining insights.

This requirement aims at enhancing the perceived usefulness of the tracked information, by preventing the possibility of getting lost among a variety of unrelated data. Providing users with correlations they are not able to see by themselves could be the key point to help them interpret their own information.

Previous studies found how experienced trackers also strive to find correlations among different data. Choe et al. (2014) showed that Quantified Selfers want to identify correlations by employing statistical tests or self-experimentation. Epstein et al. (2014) and Bentley et al. (2013) also recognised the importance of correlations among data, helping users find them through visual cuts or natural language. This requirement highlights that supporting correlation discovery is even more crucial for naïve users, as they are less goal driven (e.g. they do not have an established behaviour change goal) than experienced ones (Li, Dey, and Forlizzi 2011), and thus less motivated in actively finding these relationships. The main motivation to track, for naïve users, is curiosity towards the possibility of discovering something interesting about themselves. They see data exploration mostly as a playful activity: differently from the tracking styles identified by Rooksby et al. (2014), this way of navigating personal data is focused on enjoyment and requires an immediate engagement. Moreover, this requirement emphasises that providing clear correlations may increase the perceived utility of PI tools. Clawson et al. (2015) noted that lack of utility may motivate the abandonment of self-tracking instruments among experienced users; while Lazar et al. (2015) confirmed that a scarce perceived utility of the collected data may induce the novices to abandon their devices as well. Our study confirms these findings, suggesting highlighting data correlations to increase their perceived utility and preventing the abandonment of PI tools.

4. A new PI system

We designed and implemented a new PI system embedding the aforementioned user requirements (Figure 1). This system aims at satisfying not only the desires of the experienced self-trackers, but especially the needs of the naïve ones. In the following, we will briefly illustrate the system's key features, and how they were designed to address the requirements identified above. Then, we will describe it along its implementation details.

Our system integrates different kinds of data (e.g. sleep, heart beat rate, steps, called 'channels' in the rest of the paper) coming from commercial wearable devices and PI applications, as well as ad hoc developed

Figure 1. The system in our labs.

prototype tools. The system brings these different channels on a single flexible platform, providing a univocal entry point, repository and management interface, which allows users to add or subtract sources depending on their needs (Requirement 1).

We also designed a fruition environment to visualise all the gathered data. The display surface is represented by an 80″ screen positioned in portrait mode. The large screen is aimed at fostering engagement and immersion in the visualisations. We implemented two levels of visualisation, corresponding to two different perspectives on the same data (Figure 2), to meet the different user needs in accessing and exploring their information (Requirement 4).

The first visualisation environment is *evocative* (Evocation) and it is addressed to users who do not want to use up too much time for navigating their data. When the user stands in front of the screen, the system wakes up a data-driven reflected image of the user, which dynamically represents a particular channel of her personal data (e.g. her sleep) through a colour code (Figure 5 (left)). This representation aims at fostering the identification of the user in her data (Requirement 3). The user can move through different data channels and through time, navigating back and forth over the days, by using simple gestures. The required interaction is minimal, actualised through body movements to enhance user immersion.

The second visualisation environment is *synthetic* (Synthesis) and allows the user to get an overview of all her personal data connected with the temporal and the spatial dimension: it is addressed to individuals that want to go in-depth in the exploration (Figure 6 (left)). Different channels are simultaneously displayed on parallel timelines and it is possible to retrace the exact location in which the user was in a specific point in time. It also shows the pictures taken by the user in a specific moment and place. This representation is aimed at giving users a meaningful context through which to relive the episodes connected to their data (Requirement 2). The tools to help people interpret their data are offered as well. First, Synthesis shows the maximum and minimum peaks for each data channel, providing a perceptual cue emphasising that some data have something special, unusual or out of the average. Second, it highlights the co-occurrences between two peaks (like increase of steps per minute/increased heartbeat) in the same time unit (Figure 6(right)). These visual cues aim at helping users understand their data and find useful correlations among them (Requirement 5).

Before illustrating the architecture of the system, the data acquisition process and the data visualisation modalities, we want to stress the main differences between this system and the previous work outlined in Section 2.

Differently from the current commercial mash-up platforms (Table 1), which simply juxtapose the data coming from different sources in a unique interface, our system tries to build an integrated view of people's personal data, by providing a unique metaphor for giving access to different kinds of information (Evocation) and comparing them along the spatial and temporal dimensions (Synthesis). Moreover, many of these aggregators are focused on a single type of data (health, fitness, goals, etc.), while our system aims to provide a broader and holistic vision of wellbeing, including all the domains that constitute the individual's everyday life (e.g. social relationships, mood, location, etc.). Finally, our system is not tied to a specific ecosystem of devices

Figure 2. Visualisation environments: *Evocation* on the right and *Synthesis* on the left.

and applications, but can be easily expanded by integrating new data coming from different tools.

As for the academic works, Table 3 comes back to the PI systems and visualisations quoted in the Section 2 and compares them with our PI system against the identified requirements.

The table shows how previous works satisfy some of the defined requirements. It also highlights differences with our solution. However, none of these works addresses all the requirements together, as our system does. We claim that this is paramount to really satisfy the naïve users' needs.

4.1. System architecture

The system's architecture (Figure 3) is composed of several modules.

Input modules. External PI systems make personal data available (such as sleep or temperature) as APIs (*input data channels*). Then, a *management interface* allows the user to register the PI systems, for which she has an account, in our platform, whereas a *data gathering module* on the server collects data querying the APIs of the selected PI systems.

Processing modules. These modules run on the server and process the input data in order to provide the final output to the user. First, the *data homologation* module homogenises heterogeneous data coming from different providers, consistently with the *data model* describing each personal parameter at a *conceptual level*, and stores the data within a MongoDB database. Then, the *data analysis* module performs some statistics on the collected data in order to find meaningful patterns in the user's behaviour.

Output modules. A web server (Data Exposure) queries the database and exposes two APIs (Evocation API and Synthesis API) that can be used by the front-end web application for the visualisation.

4.2. Data acquisition

At present, we considered as input data channels the following sources, as indicated in Table 4: Jawbone Up bracelet for sleep and steps; Empatica E3 bracelet (https://www.empatica.com/) for arousal, heart beat and body temperature; Moves for locations and transport means; Forecast.io for weather; Google mail API for emails. Moreover, we developed ad hoc applications for calls, text messages, pictures taken and music listened to and ad hoc devices for tracking mood (Cena et al. 2014) and sedentary habits (the description of which is beyond the scope of this paper).

Table 3. Comparison between previous works and our system against the defined requirements.

PI systems and visualisations		
Systems/visualisations	Satisfied requirements	Differences with our system
Consolvo et al. (2008a)	• It proposes an impressionistic visualisation of the user's data (3)	• It uses the metaphor of a blooming garden to keep the individual focused on self-monitoring, but does not support identification with data. Instead, we use a representation providing an engaging experience reflecting the user's 'self'.
Medynskiy and Mynatt (2010)	• It integrates different kinds of data (1)	• It provides a set of web services that can be customised by researchers to create health applications. • It uses data related to health, whereas we include other aspects of the user's life. • It does not use ontologies for the homogenisation of data, as we do.
Li, Dey, and Forlizzi (2012)	• It uses contextual information to enrich the user's data (2)	• It provides information about events, places and people connected to specific data by using textual labels, whereas we provide contextual data visually, by using maps and pictures. This might better support the recollection of specific episodes related to the gathered data.
Bentley et al. (2013)	• It integrates different kinds of data (1) • It provides visualisations that go beyond stats and graphs (3) • It highlights data correlations (5).	• It focuses on health information, whereas we collect a wider range of data. • It provides correlations as textual messages, while we display data by using graphical representations: these can simultaneously elicit more insights than ones using natural language, which focus on specific aspects of the gathered data (e.g. 'You are happier when you sleep more'). • It does not use ontologies for the homogenisation of data, as we do.
Feltron (2014)	• It provides an impressionistic representation of personal data (1) • It provides multiple views on the same data set (3)	• It provides static visualisations, while our system gives the opportunity of interact with data. • It is focused on the author's data and users cannot visualise their own data
Baur et al. (2010)	• It supports the recollection of contextual information as well as past episodes (2)	• It provides only temporal and visual information, while our system allows users to localise data on a map, retracing their personal movements throughout the day.
Bauer et al. (2012)	• It proposes a 'passive view' that requires minimal interaction (4)	• It offers only one view on data, whereas our system provides two different visualisation modalities, allowing for exploration and analysis, satisfying different needs.
Epstein et al. (2014)	• It provides different visualisations for the same data cut (4)	• It does not connect specific visualisations with specific users' goals, whereas our system provides two different visualisation modalities, for people who have different objectives.
Huang, Tory, and Bartram (2016)	• It supports the recollection of contextual information as well as past users' episodes (2)	• It provides only a 'temporal axis' to order the collected data, while our system supports the recollection of temporal and spatial contextual information.

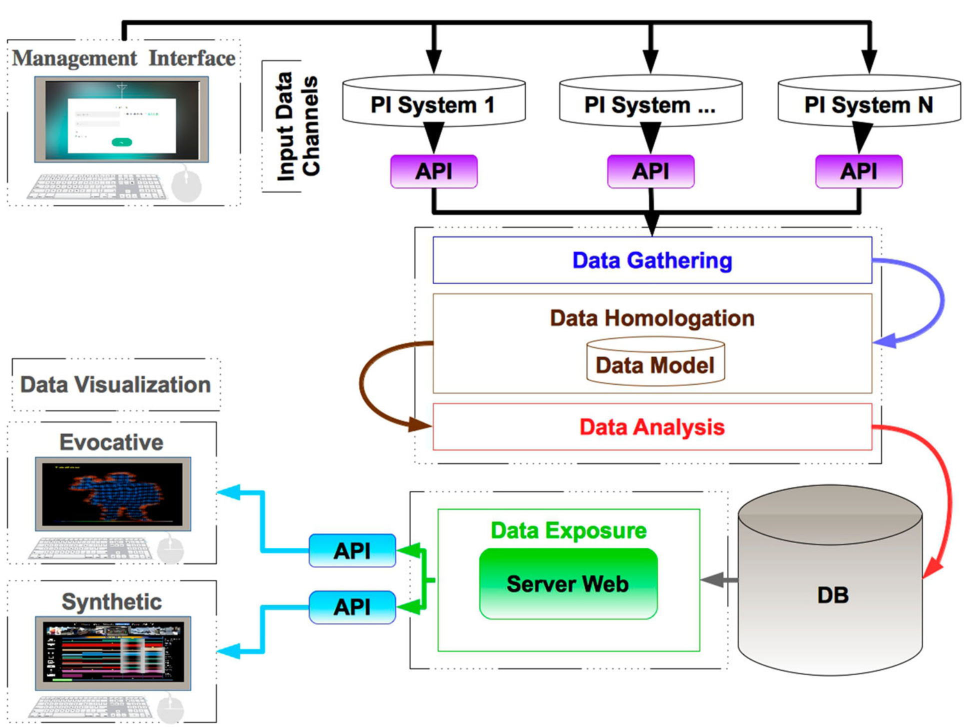

Figure 3. General architecture of our system.

Table 4. Data we used in our system.

Data	Source	Collection
Sleep	Jawbone Up	Automatically
Steps	Jawbone Up	Automatically
Arousal	Empatica E3	Automatically
Heart beat	Empatica E3	Automatically
Body temperature	Empatica E3	Automatically
Locations	Moves	Automatically
Transport means	Moves	Automatically
Weather	Forecast.io	Automatically
Email	Google mail API	Automatically
Calls and text messages	Ad hoc mobile application	Automatically
Music listened to	Ad hoc mobile application	Automatically
Pictures taken	Ad hoc mobile application	Automatically
Mood	Ad hoc tangible device	Manually
Sedentary habits	Ad hoc tangible device	Automatically

The system is flexible and other channels can be easily added. All these channels provide APIs that adopt standard technologies to access data (HTTP/JSON) and to let the users authenticate to get their data (OAuth). SSO technology provided through OAuth plays a fundamental role: we can aggregate the identities a single user presents to different PI systems by registering them once in our platform, through the management interface, which is implemented by Spring Framework. In the event that two or more providers collect data for the same channel, the user can choose the one with the highest priority.

The data homologation module homogenises data at two levels: (i) a format data type homologation (e.g. all dates are converted into UCT format) and (ii) an abstract data type homologation that maps each input data channel in the correspondent data model. Data models are represented by means of OWL ontologies and then implemented through Java classes mapped on them. We designed and developed a series of OWL ontologies aimed at modelling the different parameters of interest in the PI field (sleep, activities, time, place, etc.). For example, the Sleep Data Model provides an abstract model for the sleep channel, representing the features of that domain by means of classes (e.g. *Sleep_phase*, such as *Light_sleep* and *Deep_sleep*, *Awake_phase*) and properties (e.g. *has_total_sleep_period*, *has_total_awake_period*, *has_time_fall_asleep*). Figure 4 presents a portion of the ontology, representing the taxonomical hierarchy. If we want to integrate in our system a new PI tool that gathers sleep data, it is sufficient to create a new mapping between the tool's data and the Sleep Data Model. In this way, data models allow the independence of the channel from the provider. This is a key aspect of our system, since it allows an easy integration of new data coming from different devices, also those that are not currently available on the market, but that could be developed in the future. For example, the Sleep Data Model describes the different sleep phases.

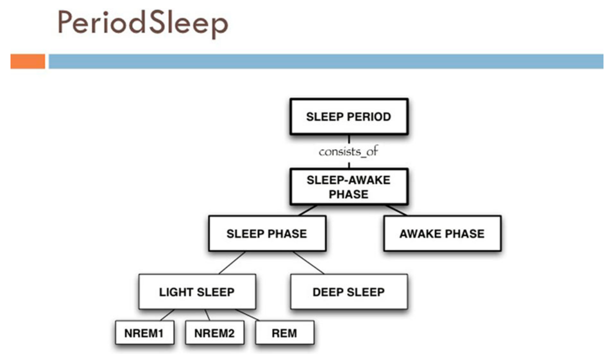

Figure 4. The hierarchy of the sleep ontology.

REM phases are examples of items that are not currently used in our system, since, till date, there are no commercial tools available that are able to recognise them. However, these phases are included in the model as they could likely be detected in the near future by new devices.

The gathered data are analysed by the data analysis module through basic statistics (mean, median, standard deviation), looking for peaks within a single channel and co-occurrences of peaks between two or more channels. These analyses could be expanded in the future with more advanced operations, through data mining techniques.

4.3. Data visualisation

Our system implements two different visualisation environments addressed at satisfying diverse user needs in navigating their personal data.

4.3.1. Evocation

Evocation is the first layer that the user encounters and it is aimed at fostering her identification in her own data: during the design phases, we were looking for visualisation designs that could be experienced as natural and immediately recognisable, thus providing a feeling of immersion and engagement. At the same time, they should not be too realistic and predictable, in order to stimulate a sense of wonder and curiosity. To this aim, we favoured more impressionistic body depictions over realistic ones, to enable an aesthetic experience, without triggering the 'uncanny valley' effect, often associated with near-realistic human representations, which can elicit a negative response from the user (Mori, MacDorman, and Kageki 2012). The result is an infovis environment where the user can interact with a joint visualisation of her body and personal data, using simple gestures.

Evocation starts with a black, empty screen. When the user stands in front of the screen, an 'apparition' animation starts; a cloud of points, coming from all directions, quickly converge, forming the user's body shape, and from this moment on it acts like a reflected image. A default channel is selected and the time point is set to 'today' (in Evocation, time points are days).

Looking at her image, the user can notice two things:

(i) The colour of the image, which is related to the value of the selected channel at the selected time point. The mapping between the channel values and colours follows a thermographic-like colour scale, visible at the bottom of the screen: white/red means a high value, while green/blue a low one. Although in data visualisation studies colour scale commonly matches the characteristic of the data and the user's task (Bergman, Rogowitz, and Treinish 1995; Brewer 1994, 1999; Borland and Taylor 2007), we used this colour code to perceptually convey a state of 'highness' or 'lowness' connecting the visualisation to the common sense idea (confirmed by psychology research as well, e.g. Ho et al. (2014)) that 'warm' colours symbolise high temperatures (and thus metaphorically high data value), whereas 'cold' colours represent low temperatures (and thus low values in data). Our main aim, in fact, was to support naïve users in intuitively understanding their data at a glance.

(ii) The halo, which is related to the averageness of a specific value: the thicker the halo is around the figure, the less common is that value for the user in her recent past. This visual cue is meant to elicit a sense of 'strangeness' with reference to a particular data channel and not to provide precise information. For example, Figure 3 represents the sleep channel: the user has slept poorly today and this value is quite unusual in her sleep patterns (the halo has a medium thickness).

Raising the left hand, the user can select a different day by moving the hand: the farther from the body the hand, the farther in the past the day (Figure 5(left)). Reaching with the right hand slightly outside of the right side of the screen, the user can call an icon menu for the selection of the channels. Using the right hand, she can interact with the icons and choose another channel (Figure 5(right)).

Both the halo and the body image do not allow for a precise quantification of the user's data: the aim of this visualisation is not to provide analytical and multiple information, but to convey a general impression about a single data channel, engaging, at the same time, the user in an immersive representation.

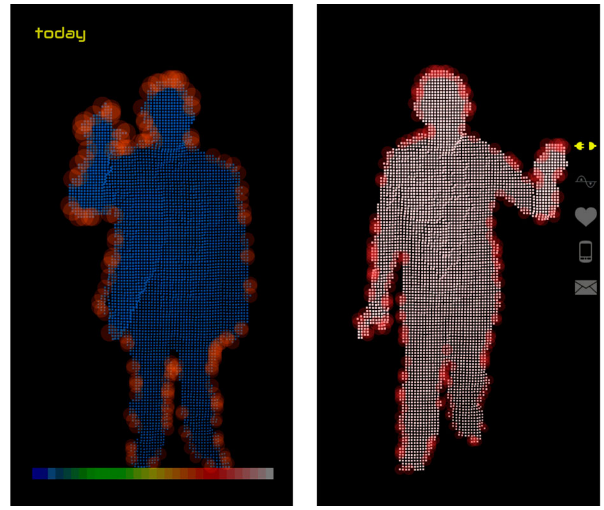

Figure 5. Day selection (left) and channel selection (right).

Finally, when the user wants to leave Evocation, she can raise both hands at the head level, keeping them in this position for three seconds. A 'dissolve' animation starts, quickly dissolving the user's body shape in a cloud of points. Immediately after, the user leaves Evocation and enters the following layer, Synthesis. The 'appearing' and 'dissolve' animations are aimed at suggesting that the user's self is somehow made of her data.

Implementation details. Evocation is composed of the following hardware components: (i) a Kinect depth sensor, for acquiring the 3D body data, limb tracking and gesture recognition; (ii) a PC, for data processing and visualisation. The framework's software components are implemented in Processing (Reas and Fry 2007).[3] Finally, the Data Exposure module, implemented by Node.js framework, gives Evocation access to the user's PI data stored in the system.

4.3.2. Synthesis

As for the user experience, passing from Evocation to Synthesis is completely seamless. Interaction in Synthesis is currently achieved through a wireless touchpad. The choice of changing interaction modality with respect to Evocation is due to the need of having a more precise and intuitive way to point, explore and analyse the collected data in a 2D environment. Although this break in the interaction modality might lead to undermine the immersion, gestures would not have allowed for such a fine-grained examination, whereas a touchscreen-based interaction would have constrained to display data on a smaller screen, whereby our initial hypothesis pointed that a larger display would have enhanced the naïve users' engagement. However, as Evocation and Synthesis are thought of as two environments

Figure 6. Different areas of Synthesis (left) peaks and co-occurrences (right).

addressed to satisfy different needs and to be used for different purposes, a lack of immersion in the transition between the two visualisations has not been considered crucial.

The touchpad is positioned on a table near the user. Synthesis starts on the current day, showing all the input data channels on parallel timelines: the user can zoom the time in (down to one hour) or out (up to weeks, months and years). The view is divided into five logical areas (Figure 6(left)): (A) time unit selector; (B) pictures related to the time unit; (C) timelines for the connected Input data channels; (D) time/space connection cursor; and (E) spatial map. Positioning the pointer over any element of the chart, Synthesis displays a tooltip with more details about this element (e.g. the heart rate value for that instant).

By swiping on the touchpad the user switches to the preceding or following time unit (e.g. previous/next day), depending on the swipe direction. By dragging the time/space connector cursor (D), the user makes the spatial map (E) slide in relation to the progress of time in the timelines (C). In this way, it is possible to see where the user was at a specific time point related to particular data values. The pictures, positioned at the top of Synthesis (B), flow along with the proceeding of time as well. The spatial position on the map connected with the user's pictures is aimed at eliciting her memory, triggering the reminiscence of her personal experience related to the tracked data.

Synthesis also provides perceptual cues aimed at highlighting unusual or special values, in order to help users interpret their data. Peaks within each input data channel are shown by means of 'triangles' (highest values) and 'diamonds' (lowest values), while the co-occurrences of peaks among different input data channels, in the same time unit, are displayed as glass tubes that vertically embrace diverse timelines (Figure 6(right)). Triangles and diamonds have been chosen due to their different shapes, in order to visually mark the diversity between high and low values, without conveying any positive or negative meaning.

Implementation details. To allow users to seamlessly move from Evocation to Synthesis, a communication scheme using a websocket connection has been used. When the user exits from Evocation, it sends Synthesis one token. Shortly after, Evocation hides its window after the final black frame of the animation. Nearly at the same time, Synthesis, alerted by the token, starts a 'fading in' transition from an identical black frame to its visualisation layout. Synthesis is implemented in Javascript and runs in a browser. Data rendering is achieved through D3.

5. Evaluation

5.1. Procedure

We recruited 24 participants through mailing lists and snowball sampling (age $M = 32.7$, $SD = 10.1$, females = 13). We used three mailing lists containing a total of 430 members. Initially, 42 individuals were interested in our research. Sample selection followed a purposeful sampling method (Gobo 2004; Marshall 1996), looking for individuals showing curiosity and interest in using a tracker. We ascertained this aspect in the same manner we did in the diary study. We screened participants through a telephonic interview (see Appendix A.2 for the screening criteria). Although we cannot claim the representativeness of our sample with reference to the demographics of the entire population (e.g. all our participants were Italians), we balanced it with reference to age and profession in order to increase its heterogeneity (Marshall 1996). The decision of settling for 24 participants came when we realised that additional data would not have produced substantial new findings for the goals of our study, following a data saturation criterion (Bowen 2008).

The primary objective of the study was to assess whether the five user requirements embedded in our system could really satisfy the users' needs: this would provide other researchers with reliable reference points to be considered in PI design. To this aim, we decided to employ qualitative methods: we wanted to gather rich insights, even unexpected, about how naïve and experienced self-trackers were impacted by the five requirements, rather than to assess our solution with respect to precise quantitative variables. For this, we conducted semi-structured interviews instead of employing questionnaires or Likert scales. As a secondary goal, we preliminarily explored whether the system could affect users' intention towards change. To this aim, we asked individuals in the screening interview about their orientation towards behaviour change, following the TTM. We based the screening questions about this aspect on Etter and Sutton's (2002) work.

The majority of participants (U1–U16) had the same characteristics of those involved in the diary study (naïve users): they did not have any previous experience with self-tracking, but were curious or interested in this technology. They did not have any specific need to track a specific behaviour. Therefore, curiosity towards the possibility of discovering something about themselves was their main motivation to delve into tracking, as highlighted in the diary study, which explored the reasons for collecting data of this kind of users. In addition, we included eight participants (U17–U24) who regularly used self-tracking devices (experienced users) to assess whether the requirements embedded in our system could satisfy their needs too.

As to behaviour change stages with respect to specific health behaviours, participants with no experience in PI were in the precontemplation stage (8 out of 16), contemplation stage (6 out of 16) or in the preparation stage (2 out of 16).

Those in the first stage were precontemplative about all the health behaviours that they could track in the study (eating habits, physical activity levels, sleep, stress), since they did not even think about changing any of them; participants in the second stage were contemplative about eating habits (1), physical activity level (4) and stress (1); those in the third stage were in preparation with regard to the physical activity level (2). Experienced participants, instead, were in the preparation stage (1 out of 8), action stage (3 out of 8) and in the maintenance stage (4 out of 8) with respect to a specific behaviour. The participant (1) in the preparation stage was in preparation with regard to physical activity. Those in the action stage were dieting (1) and running (2), while participants in the maintenance stage were regularly exercising (3) or trying to maintain a certain amount of sleep (1) (see Table 5). For almost all the experienced users involved in this study, it is safe to say, the main reason to track was to change, or maintain, a certain behaviour. An exception is represented by the experienced user in the preparation stage who mainly tracked for documenting his activities.

The majority of participants were well educated and relatively affluent. Four participants stopped studying at middle school. Five participants held a high school diploma, 10 a bachelor's degree, and 5 a master's degree.

Participants' ages ranged from 18 to 58. They were undergraduate students (2), PhD students (2), post-doctoral researcher (1), university professor (1), psychologists (2), employees in an ICT company (2), workmen (3), lawyer (1), physicians (2), designers (2), software developers (2), insurance agents (2) and executives (2).

Each participant had to use a wearable device and/or a set of PI apps: U1–U4 and U17–U18 used an Empatica E3 bracelet for arousal, heart beat and body temperature, Moves for locations, an ad hoc developed app for calls, MyFitnessPal for food; U5–U11 and U19–U21 used a Jawbone Up bracelet for sleep, steps, mood and food, Moves for locations and an ad hoc developed app for pictures; U12–U16 and U22–U24 used Moves for locations, Google mail API for emails, Expereal for mood and two ad hoc apps for music and pictures. Devices and apps were assigned depending on the specific preferences of the participants, with regard to the parameters to be tracked.

Participants had to wear the devices and use the apps for one month and a half. For the automatically tracked data, participants tracked them almost every day (except U4, U6 and U20 who often forgot to wear the bracelet). For the self-reported data, they experienced the same issues as those participating in the diary study (e.g. failing to report due to lack of motivation). However, the compliance in accomplishing this task was sufficient to gather a continuous flux of data to be visualised in our

Table 5. Relationship between stages of change and specific behaviour in the sample.

Stage of change	Eating habits	Physical activity level	Stress	Sleep
Precontemplation	8 Inexperienced (U3, U4, U5, U6, U8, U12, U13, U14)			
Contemplation	1 Inexperienced (U7)	4 Inexperienced (U9, U11, U15, U16)	1 Inexperienced (U1)	//
Preparation	//	2 Inexperienced (U2, U10) 1 Experienced (U20)	//	//
Action	1 Experienced (U18)	2 Experienced (U23, U24)	//	//
Maintenance	//	3 Experienced (U17, U21, U22)	//	1 Experienced (U19)

system: participants self-reported data at least one time every two days (except U23 that stopped self-reporting after 10 days). Participants maintained this regularity even in the last part of the trial (between the second and the third interviews). One participant (U15) dropped out of the study after 9 days due to personal reasons. Participants did not have to keep a diary as in the previous study. We inferred their compliance in tracking their data from the information gathered and the answers given during the interviews.

Each tracker's participation included three interviews. During the first, we installed the applications on their personal phones, set up the devices to be used and asked about their initial expectations of our system. After three weeks, we invited them in our labs where the gathered data were imported in our system; then, we presented our solution and gathered feedback. The researcher first gave a demonstration of the system. Participants stood and walked in front of the screen in the sensing range of the Kinect. They could use the system freely for as long as they liked, interacting through gestures with Evocation and using a touchpad with Synthesis. The average period of engagement was 30 minutes. Participants gave feedback on our system in a talk-aloud format, speaking freely as they examined its features. No specific tasks were given and they were observed by the researcher. After the session they were interviewed for 30 minutes. Interviews were qualitative in nature and they were audio recorded (see Appendix A.2 for the questions asked). Finally, after three more weeks they came again to our labs to examine one more time the data they gathered. The set-up was identical to that of the second interview, but participants could explore a more substantial data set, given the longer time for data gathering.

Results from the interviews and the talking-aloud sessions were first analysed through a thematic analysis, like in the diary study. Analysis followed open and axial coding techniques (Strauss and Corbin 1990). Results were coded independently by the first and the last authors by taking apart sentences and labelling them with a name. Then, findings were reviewed segment-by-segment to evaluate consistency in the application of codes, then resolving all inconsistencies. Results were finally grouped along with the requirements found in the diary study.

5.2. Results

5.2.1. Acceptability and usability issues

Reactions to the system were positive. Both experienced trackers and naïve users valued the system and thought that it helped them identify unusual and surprising patterns in their data, encouraging them to continue their exploration.

Participants perceived the gesture-based interaction with Evocation as simple, enjoyable and playful. U5, for example, reported that 'it's like playing with the Wii'. However, some participants, like U22, found some gestures unnatural: 'it doesn't make sense that I have to raise my hands in the air to change visualization. Why have I to do that?'. This suggests that some gesture-based commands should be rethought in order to meet users' expectations. The main issue that participants encountered in interacting with the system was related to the different interaction modalities of the different visualisation environments. Participants were disoriented in using a touchpad for interacting with Synthesis after having used gestures to act in Evocation. Many participants wished an extension of the gesture-based interaction even in Synthesis.

Another problem was connected to the metaphor employed in Evocation. Although the majority of participants agreed that the body colour and the halo provided an intuitive code for interpreting the data at a glance, some of them (8 out of 21) questioned the understandability of the halo in providing information about the averageness of their data: they stated that there was no 'natural connection between the thickness and the regularity of those data', as said by U8, while in the colour code they found a more immediate relation between the grade of the values and the colour scale.

Another point to be noted is connected with the suitability of the solution for the home environment and its integration in the users' everyday life. U18 noted how this system could be put in her home: 'It's like a piece of furniture, it's beautiful to see … it could be always turned on, a pretty piece of furniture', she said, even if it should be put in landscape mode to find 'a collocation in my home', as said by U19. However, this enthusiasm

Table 6. Summary of the study's key findings.

Requirements	Findings	Differences between naïve and experienced participants
Requirement 1	A unique entry point for all the data increases the sense of control. Flexibility in adding and removing PI tools is a key feature for addressing users' changing needs over time.	Experienced participants are more aware of the critical issues related to data storage (e.g. ownership of data) than naïve ones.
Requirement 2	Pictures are useful for reliving episodes far in the past and particular occasions. Locations were confirmed as fundamental elements for the reminiscence of everyday episodes associated with particular data.	All participants found that remembering is a means for understanding the reasons underlying a particular data pattern and a tool to enrich data with meaningful lived episodes.
Requirement 3	Evocation is capable of enhancing users' identification with their data. Experienced users find Evocation immersive but not so useful.	Experienced users have an instrumental perspective on data, while naïve ones prefer to be surprised by immediate representations.
Requirement 4	Evocation better addresses the naïve users' needs, whereas Synthesis better satisfies the experienced users' interaction style. The proposed visualisations were not able to accommodate all the users' exploration styles.	Whereas naïve users are 'passive' in exploring their data, experienced ones are 'active' and want to manipulate them with more fine-grained tools.
Requirement 5	Correlation highlights increase the value of data.	Naïve users claim a design capable of subtracting, rather than adding, elements from the visualisations, while experienced ones focus on deeper exploration of data.
Behaviour change	The system is able to raise awareness of users that are in the TTM's precontemplation and contemplation stages. It does not satisfy the experienced participants' needs with reference to behaviour change.	Naïve users express the need for receiving recommendations about goals to pursue and ways to achieve them, while experienced users want to set their own goals and seeing progression towards them.
Sociality	The system lacks social features.	All users highlight that social exchange of data is connected with close interpersonal relationships.

was limited to the participants that had some kind of previous experience in self-tracking. The majority of naïve users found the system too invasive and cumbersome: they wished for a solution suitable to be displayed on the devices they already owned, like a television or a desktop pc.

Despite these criticalities, the study showed that the system is able to satisfy the requirements found during our diary study as well as the validity of the requirements themselves. Table 6 gives a snapshot of the key findings of the study, which will be extensively reported in the following.

5.2.2. Requirement 1: Promote the integration of different sources of data

All participants appreciated the availability of all their data in a single location, where to connect, disconnect and eventually reconnect different devices. They highlighted that having a unique entry point for a variety of data gave them more perceived control upon them, increasing the perception that they were in 'a private and secure place'. Control, in fact, seemed to be one of their main concerns: U4, for example, stated that

> maybe I could use a tool for a while and then choose to switch to another one. All those data would be lost, closed in a unreachable box … here it seems that I can have them all under my control, all my present and past information.

Experienced participants (6 out of 8) further stressed the need to flexibly integrate different devices over time, not only for their rapid technical obsolescence, but also for possible changes in the tracking needs. U23, for example, noted how wearable devices and mobile apps change rapidly, as the technical improvements in terms of accuracy, unobtrusiveness and battery life are impressive from year to year, and a user should not start over and over again to recollect her data: 'we should be free to change our devices without worrying to restart collecting our data'. U19, instead, emphasised that

> Now I'm tracking my sleep because I think that it's important, but what about two years from now … who knows what I will need to track at that point … maybe I will want to track my stress … and maybe after a year from then I will have again to track my sleep. All the data that I'm collecting now should then be preserved.

Three of the experienced self-trackers further highlighted that giving all the data to a unique provider could also make them more vulnerable: 'this system could know everything about me, and use my data for its own purposes, for example selling them to someone else. Who knows?', said U20, even if, after a while he corrected himself stating that 'however, a lot of services already own my data, maybe I'm worrying too much. Facebook knows a lot of things about me, but I usually don't care so much'. This supports our prior belief that experienced self-trackers might have a more critical view on the complex problems that lie behind the storage of personal data.

All participants (except one) also positively judged the possibility of adding 'new channels' to the system, without changing the visualisation modality. This, however, is not exempt from design challenges. U1, for example, was enthusiastic for the possibility of tracking an 'internal' parameter like stress, noting that 'Maybe soon it will be possible to track also something that pertains to the mind.' This wish raises the interesting question about how much flexibility, with regard to interface design, our system is capable of supporting: could specific kinds of data (e.g. mental states, complex activities) be easily integrated in our system? Or could they have peculiar 'shapes' that will not easily fit in our designs? Another question is related to the number of channels that our two visualisations are able to support without jeopardising their understandability and intuitiveness: this points to the need of reflecting on how much flexibility is desirable as well as sustainable.

5.2.3. Requirement 2: Support users in remembering their data

By looking at Synthesis, all participants experienced moments of discovery, learning something about their behaviours they were previously unaware of. Synthesis was capable of eliciting richer experiences than those provided by the tools used in the study, as participants were able to recollect the memories associated with specific patterns of their data. Most of them (17 out of 24) reported how this was a way to make them more valuable, as each single gathered information was perceived in its uniqueness.

U5, for example, reported that in the Jawbone Up application

> I could not figure out why I slept so poorly in certain days … here I can easily see that I was on business trips … . Even if they were close to my work place, I think that they stress me out a lot … this is likely why I didn't sleep well in those days.

U3 remembered why on Saturday evening his heart beat rate was so high: 'I was just running to catch the bus, my heart was about to explode … I almost forgot that before', he said, looking at the spatial map indicating his current position. Remembering, then, was considered by participants a means to understand the reasons underlying a particular data pattern, as well as a tool for transforming neutral information into meaningful lived episodes.

Pictures were judged more useful in reliving episodes far in the past or particular occasions, such as vacations, than in recollecting daily events:

> it's unusual to take photos of the places that we are used to visit during our daily routines, – said U14 – but they can add a great value for the data collected during the holydays, or when we are out for a dinner … look at this, I was at the sea, I took it just before going to the restaurant … this is why I was in such a good mood … these pictures can fix some important events around which mentally organizing the data, I think.

Pictures, then, work as landmarks that might support users in anchoring their data to specific events, helping them orient themselves throughout the collected information, as well as enriching its temporal organisation with supplementary meanings.

Locations, instead, were confirmed as fundamental elements for the reminiscence of everyday episodes associated with particular data. U3, for example, emphasised how

> I can reconstruct what I've done throughout all the week seeing my position and my movements on the map … And I can also recollect why I was calling a person … This call for example to my wife … I had just left from my office, I called her only to tell her that I was arriving.

Several participants (5 out of 24) stressed the importance to know 'the people that were with me in a particular place', as said by U11, to have a richer context through which to re-experience their data.

5.2.4. Requirement 3: Support users in identifying with their data

Evocation was considered an intuitive visualisation of the user data and the system was also capable of simplifying their management, exploration and interpretation. In fact, several participants (7 out of 24) thought that it was similar to watching a sort of 'personal movie'. Participants (20 out of 24) were fascinated by their reflected image. They considered this representation as extremely 'perceptual', not requiring difficult interpretation processes, but immediate as 'watching a movie, when you only have to look at, and not so much to think about … because everything is already clear', as said by U9. The reference to the movie experience well outlines how a sort of 'passive' fruition of the collected data was perceived as valuable by the naïve users, provided that it can be an engaging and enjoyable experience. U10, for example, stressed that exploring data should not be a work:

> I don't want to waste my time in analyzing myriads of data … I like the fact that I don't have to learn to use this system, that I can simply look at myself [i.e. the data-driven reflected image] and see whether the data are right or wrong, better or worse than yesterday. It's somehow fun.

Naïve users, therefore, are more oriented in navigating data for the mere pleasure of discovering something unexpected, seeking the engagement of the experience, rather than the usefulness of the analysis.

Nevertheless, experienced self-trackers (5 out of 8) were doubtful about the usefulness of Evocation: U17, for example, reported that 'This is not really surprising for me, I think that it's not useful for finding insights in my data. It's a funny game, but I can't reach my goals with this', confirming that experienced users prefer numbers and graphs to make their data actionable. This shows how experienced users who track to change their behaviour have a more instrumental perspective on data: they are less inclined to casual exploration since they perceive the data examination as a means to a further end.

Participants' comments also endorsed the idea that Evocation is capable of enhancing their identification with their data. Participants (19 out of 24) explained how they perceived their body shapes merged with their data as more personal and closer to their selves. For U4, for example, 'it's me over there, or better, me in a specific state … or behavior', while U6 particularly stressed the difference between the Jawbone app's graphs and Evocation:

> there [in the Jawbone's app] you can look at your data from a certain distance, they are transformed in quantities, in bars … here I think that it is different … it is difficult to ignore that they refer to you, that they belong to you.

Seeing a reflected image moving when the user's body was moving was considered as immersive, giving the 'sensation of being there', U8 said. This sense of projection was shared also by the majority of the experienced self-trackers (6 out of 8), confirming that Evocation is effective in making the data 'more personal', even if not all participants believe it to be capable of making the data 'more actionable'.

5.2.5. Requirement 4: Offer different views on data

The majority of participants (20 out of 24) commented on how much they appreciated the possibility of switching between two diverse visualisation modalities. While Evocation was perceived as an intuitive view that could be always kept active, Synthesis was thought of as a 'deeper' way to look into personal information, requiring a higher attention even if 'not very demanding' as noted by U4. U14 specified that

> even if it's not difficult to learn how to interpret these data [Synthesis], I think that the first one [Evocation] for me is more than enough. I need only to glance at it for understanding what happened to me today.

U18, instead, highlighted the importance of the Synthesis:

> in the other one [Evocation] I cannot easily compare different data types. I think that seeing how these data vary together over time is the real value of this system … Here I can clearly see that my stress level raises when my steps and calories are not well balanced;

while U17 reported that 'in the second one [Synthesis] I can check every single information, visualizing its numeric value … I think that it's important when you have a specific goal, such as making a minimum amount of steps every day'. To summarise, almost all the naïve users (except U2) wished for a 'passive' experience when interacting with their data, while experienced participants asked for a more active role.

However, results showed that the modalities proposed were not able to accommodate all the users' exploration styles and further visualisations are likely needed to satisfy them all.

> It should be possible to compare only specific data … It would be interesting to be able to ask the system what happens to my sleep when I'm in the mountains? But it should be direct as making a question and receiving an answer,

noted U19, while U21 stressed how it could be difficult to verify hypotheses about the factors that may affect her behaviour with these current visualisations. These findings highlight that the current visualisations might not sufficiently support self-experimentation and may not completely satisfy experienced users.

5.2.6. Requirement 5: Highlight data correlations

All participants appreciated the system's ability to highlight the peaks related to their data and the co-occurrences among different channels, stating that these features simplified the interpretation of their information. U24, for example, reported that 'these triangles and diamonds make me immediately aware that something has happened in my data'; while U6 further emphasised how 'they focus your attention where presumably there is something of important, otherwise it could be easy to get lost'.

However, several participants, like U3, said that 'The peaks in the heart beat is the only data that I need. The other ones are only noise', while U8 underlined that 'I think that we don't need to have all these data. A lot of information is always the same … I mean, it's like a baseline that I don't need to see every time.' This shows how some data displayed in Synthesis were perceived by naïve users as not crucial, and could then be hidden to simplify the view even more. They claimed a design capable of subtracting, rather than adding, elements from the visualisations, even at the cost of narrowing the perspective on the factors that may affect a particular parameter. U1, for example, reported that 'I should be allowed to hide some parameters to make the visualization cleaner

… for example, the body temperature, I don't think that it's a key aspect to see for me.'

Four participants also expressed the desire to see more intelligence in the system, suggesting that the provided aids were insufficient in finding useful correlations for their everyday life: U9 said that 'you need to have suggestions on how to manage these correlations. Which of them are really important for my daily life?'; while U11 noted that 'these two peaks are somehow something risky for me? […] It should be important to be advised on this point'. Correlations should then be paired with suggestions about their relevance for the specific individual.

5.3. Discussion

On the one hand, the above results show how our system can satisfy the requirements found in the initial diary study, although some improvements are still needed. On the other hand, such findings confirm the usefulness of the requirements themselves, as integration, reminiscence, identification, variety and correlation were considered important features by the participants to effectively explore and make the collected data actionable. The satisfaction of these requirements made users feel closer to their data, enhancing the perception of their value. In what follows, we go back to such requirements, discussing points we consider relevant for the PI discourse.

5.3.1. Requirement 1: Promote the integration of different sources of data

Users showed appreciation for the possibility of integrating different data sources in a single platform. Our results confirm findings emphasised in previous research, where it has been noted that users search for tools capable of integrating diverse information (Bentley et al. 2013; Choe et al. 2014; Li, Dey, and Forlizzi 2010; Liu, Ploderer, and Hoang 2015; Whooley, Ploderer, and Gray 2014). The initial requirement pointed to the importance of having a unique 'actor' in charge of guarding users' data, for privacy and security reasons too. Our study results, however, present a more nuanced picture. If, on the one side, having a single entry point for all the collected data may represent a 'comfort', as it avoids the loss of precious information, as well as a 'guarantee', in terms of increased control over data, on the other side, it may be risky to give excessive 'power' to the data owner. The knowledge that PI systems might have of their users has been frequently described as 'creepy' in previous research (Lupton 2016; Tene and Polonetsky 2013). Warshaw et al. (2015), for instance, found that users perceive artificial entities capable of knowing their inner traits precisely as creepy. Our study points to the fact that if naïve users may not be aware of this threat, for some experienced users this may become a serious concern. Therefore, the mere integration of data in a single platform could not be sufficient to satisfy all the needs of privacy and control. Designers should focus on increasing the user's perception of the system's trustworthiness, by giving her the possibility of knowing all the data it keeps, what kind of use it makes of them and all the third parties with which it could share them. Moreover, they should give users the possibility of deciding what kind of information to preserve, modify or delete.

5.3.2. Requirement 2: Support users in remembering their data

Remembering the context in which data were collected showed an increase in the value of the data themselves. Findings confirmed the importance of spatial cues in triggering memories. Moreover, they assessed the positive role of pictures to remember the past, especially when associated with important moments of the user's life. This might entail the need for finding ways for supporting data curation, in order to motivate the user to select and preserve images that are connected with significant episodes of her past. Elsden, Kirk, and Durrant (2016) precisely emphasised the importance of curation when personal data are used to remember the past. Moreover, several participants also suggested to consider the people met during a day to expand the context of data. This points to the notion of 'enriched' context that should be probably addressed in the future, when a wider variety of data will be easily available. Context-aware research referred to context as both the external (e.g. features of the environment, location) and internal (e.g. user's goals, tasks, etc.) factors that may affect the user in a specific moment (Prekop and Burnett 2003). Our research highlighted the importance of location, but PI should also explore other environmental elements, as well as 'internal states', that could enhance the contextualisation of the gathered information.

5.3.2. Requirement 3: Support users in identifying with their data

This requirement pointed to the need for providing intuitive forms of data representation in order to foster user identification with them. While our findings confirmed that Evocation fulfils the requirement for naïve users, they also highlighted that it could not fit all the needs of the others. This is due to the diverse tracking styles that characterise them: if naïve users prefer an engaging and playful mode of data exploration, our experienced participants (who mostly tracked for behaviour change) have a more instrumental perspective on their collected information. However, it does not appear that they devalue the possibility of identifying with their data at all, i.e. questioning the validity of this requirement per se. Instead, they seem to put in doubt the capability of Evocation to support the achievement of their behaviour change goals. It could be then interesting to explore whether Evocation could be also made appealing for people who are trying to change their behaviour. This could be done by making the Evocation more 'actionable', e.g. by allowing users to set the 'meaning' of its colour codes and the halo. For instance, users could customise this view by putting the halo in relation not to the averageness of a specific value, but to the distance needed to fill the gap for their daily objective (e.g. the thicker the halo around the figure, the more the steps to reach the user's daily quota). This could turn the Evocation mode into a sort of 'glanceable' visualisation. Glanceable feedback, i.e. brief sessions of interaction with data in the form of glances, has been considered important when users interact with physical activity trackers to monitor and regulate immediate behaviour (Gouveia et al. 2016). By providing some useful information in a glance, Evocation could then support the need of having actionable data related to the user's goals, also supporting her identification with data.

5.3.3. Requirement 4: Offer different views on data

This requirement stressed the need to design at least two different visualisation modalities, one more direct and intuitive, the other one more detailed and comprehensive, in order to fulfil the needs of different kinds of user. However, while findings confirmed that such views satisfy the desires of a large quota of users, they also emphasise that the proposed visualisations are not capable of accommodating all the diverse users' exploration styles. Therefore, it appears that a supplementary view on data is needed, in order to allow users to compare specific data sets, and test specific hypotheses. To this aim, Karkar et al. (2015) suggested supporting people in defining and testing hypotheses, by automating experiment design and proposing customised study plans based on a person's selection of study duration. Our findings stress that users should be provided with an interactive space in which to freely manipulate different variables, or be allowed to compare particular data sets, by formulating specific questions. On this point we think that further research is required. However, this redefines the requirement, by recommending to design at least three visualisation modalities when the aim is to promote the spreading of PI tools to a larger audience.

5.3.4. Requirement 5: Highlight data correlations

This requirement emphasised the importance of providing users with correlations they are not able to see by themselves, in order to help them interpret their own data. Our findings confirmed that correlations are fundamental to increase the perceived value of the collected information. However, some participants wanted reduction and focus on specific parameters, instead of comprehensive takes on the entire corpus of data. They also demanded a design capable of subtracting, rather than adding, elements from the visualisations, even at the cost of narrowing their vision. This suggests that focused cuts, as those proposed by Epstein et al. (2014), could be valuable for those users that are interested only in specific parameters. Observations that point to punctual correlations, also, as those given by Bentley et al.'s Health Mashups, could be useful for those users driven by the interest in connecting specific parameters, but who do not want to be overwhelmed by a plethora of co-occurrences. Such users might be represented by both people just curious about their data, who want to keep their visualisation as simple and 'clean' as possible, and individuals with a specific behaviour change goal, who are likely to be selectively interested only in those particular parameters that may affect their target behaviour: this highlights that there might be cases in which we need to design for more 'frugal' and focused visualisations.

5.4. Design considerations

We now discuss other findings from our interviews. The first point is related to the capability of the system in triggering behaviour change processes, while the second one is connected with its social aspects. Then, we will outline three further considerations for design.

5.4.1. Behaviour change

When reflecting on their data through the provided visualisations, participants found opportunities to change their behaviour. Their experiences show how visualisations can help users, especially at initial stages of behaviour change.

On the one hand, naïve users, who were mostly in the precontemplation and contemplation stage with respect to specific health behaviours (Prochaska and Velicer 1997), found benefits in understanding baselines and anomalies in their habits, making them reflect on opportunities for change. U3, for example, stated:

> It's so surprising to see that when I don't use the phone just before going to bed I sleep so well. This is something

> that I'll have to keep in mind, because it would be an extremely easy thing to change in my life.

Participants (10 out of 16) became aware of behaviours that they were previously unable to see with the assigned tools, raising their consciousness about the facts that may affect a problematic behaviour (Mori, MacDorman, and Kageki 2012). U4, for example, stated that he did not receive any value from tracking his data by means of the tools he used. However, when he saw his data visualised in the system he said 'Seeing all these data together changed somehow my perspective on this kind of tools. There are some patterns here that I really need to change. My stress is too much high when I don't walk enough.' However, naïve users expected more proactivity from the system in recommending the behaviour to be changed. U6, for example, said 'I would need suggestions based on the gathered data, like I know that your heartbeat rate raises when you call him, it's better than you don't drink coffee first!'; while U10 emphasised that 'having all my data means knowing what is better for me, what should I maintain or change … This system should do this for me, recommending what behavior I should change and how'.

On the other hand, experienced self-trackers (5 out of 8) who already had a particular behaviour change goal stressed the impossibility of setting their own objectives and seeing progression towards them: their behaviour change needs were far away from those of naïve users. U17, for example, reported that

> Certainly, having something that highlights whether I reached my daily quota or not would be helpful for me … This could also be done for longer periods of time … for example I may have a minimum amount of steps that I want to do every week, and it would be great if the system could display my progress toward it on demand.

A similar request was made by U21 who has been running for one year to maintain a certain level of physical activity in her daily routine: 'I'd like to display the impact that a diminished weekly activity would have on my health … on my blood pressure, sleep, weight, for example … maybe it could help me maintain my running habits.'

Naïve users appreciated highlights on data aimed at making them aware of behaviours that were potentially problematic or harmful for their wellbeing. They also demanded recommendations about the actions to be taken to address the highlighted issues, stating their inadequacy in identifying what could be right for themselves. Experienced trackers, instead, wished for tools capable of supporting their actions in pursuing a certain behaviour change goal, by allowing them to explore in their consequences changes in their daily behaviour.

5.4.2. Sociality

All the participants stressed the private nature of their data. However, they highlighted the lack of a sharing feature in the system. This was not meant for posting personal information on social network platforms, such as Facebook or Twitter, where participants, as U3 reminded, could lose their control on them; rather, they wished to exchange and compare information with their significant others. U8, for example, noted how 'this system is too much focused on the single individual. I'd want also to see the data of my friends and let them see mine'; U21 wished for a feature that could allow her to match her data with those of her running companions; while U17 stated that 'I'd like to match my sleep patterns with those of my wife … I'd like to see if there's something that affects the sleep of both of us'. Fifteen participants felt data were potentially shareable. However, this social exchange was mainly connected with close interpersonal relationships, as a way to show something important to other specific individuals, or to share significant information among a strict circle of users. For example, participants highlighted how saving specific data sets, connected with a particular time period or experience, could be considered as a means not only for preserving a memory, but also for sharing it. U20 stated how 'my data connected with a photo of the place in which they were gathered could become like a postcard, to send to my friends, to say them that something had happened in that moment'. We think small packs of data referring to a specific time and a specific place are more likely to be shared than lifelogs, and they could function as a digital memento to be exchanged with a selected number of individuals.

Given these findings, we now present three further opportunities for designers, which can be added to the requirements identified above, in order to orient the design of our future work and other future PI systems. However, such considerations are only grounded in the users' explicitly expressed desires, which emerged during the evaluation study. Therefore, since they were neither implemented in the system nor tested on the field, following Hekler et al. (2013) we consider them as design hypotheses, which will require additional testing to fully prove their validity.

5.4.3. Design for progression

Users with different experiences in tracking and visualising their data, as well as in using data for influencing their own behaviour, have different needs when interacting with a PI system. Experienced self-trackers who have a behaviour change goal may need support for proceeding in self-experimentation, as their motivation to track can be driven by a desire to understand the facts that may affect their behaviour: these functionalities are far away from the needs of naïve users, who first have to make sense of what is happening on the screen and not to get lost in a huge amount of quantitative data. However, several naïve users changed their attitude towards the system between the second and the third interviews. U1 stated that

> The first time I was only curious to see what happened to me in the two weeks before, and there were too many things on this screen … Now I understood more or less how all this stuff works, and I'd like to ask the system some specific questions.

This shows how the users' needs may change over time, especially when their abilities of understanding how to make their data actionable increases. PI interfaces should make more complex tools for acting on data available as long as the users' skills in exploring and understanding them improve, and as long as specific behaviour change goals emerge. Previous works also highlighted the relevance of supporting users during the visualisations of their data. Bentley et al. (2013) used natural language to highlight correlations among information in order to support trackers' behaviour change, whereas Karkar et al.'s (2015) tool provided aids to formulate hypotheses about a certain behaviour they want to test. Epstein et al. (2014), instead, emphasised that different trackers may appreciate different visualisation modalities. Here, we confirm the importance of satisfying diverse visualisation preferences, but we further propose to progressively adapt visualisations to the users' levels of expertise: we point out that these needs may vary not only from individual to individual, but also from time to time in the same individual. Starting with simple visualisation modalities and progressively increasing the number and complexity of the tools available could avoid the initial bewilderment given by a set of overabundant functionalities, tunnelling the user's experience on a path that adapts itself to her changing needs.

5.4.4. Design for personalisation

The majority of experienced self-trackers consider a PI system a sort of personal trainer, which can work together with them to formulate hypotheses and build interventions tailored on their goals. This should be focused mostly on long-term plans: from their perspective, to correctly set them, a wide range of variables needs to be considered, variables that could hardly be taken into account by a human, but that could be computed by an intelligent system. U21, for example, stressed that 'Having all these data should allow this system to provide me with personalized plans that may work until I reached my goal.'

Without having established objectives in mind, instead, naïve users wished for recommendations shaped on their past behaviours, which could help them reflect on what is better for them. U7 explained:

> I don't know what could be useful for me at present. Maybe eight thousand steps are sufficient to be well today because I haven't eaten so much … The system should advise me from day to day on the basis of my data, but I'm not interested in following a diet or a training program … nor to invest efforts in telling the system what to do … it should know it.

These users, then, are more focused on short-term matters, seeking insights that could drive them accordingly to their current situation.

These results suggest that personalisation could help users make their data actionable. Previous research also recommended the employment of personalised features in PI. Schmidt et al. (2015), for example, proposed to support individual fitness goal achievement by combining goal data with performance data, in order to provide tailored plans based on the collected training information. Lee et al. (2015), instead, employed a reflective strategy that prompts people to reflect on their priorities and goals to encourage them to customise their fitness plan by themselves.

On the one hand, our findings point to the usefulness of tailored plans for experienced users with a behaviour change goal, who have the willingness to work with the system to set interventions and realise them. Such users are focused on the distant future: PI systems should then offer opportunities for achieving their long-term goals through personalised plans. On the other hand, findings point out that naïve users are more oriented to the present and to short-term issues that may not be necessarily related to behaviour change. They appear to be less inclined to spend time 'collaborating' with the tool to discover what could be most suitable for them: personalisation, here, should then be addressed at providing just-in-time advises supporting them in their daily life, without requiring strong efforts in terms of usage settings.

5.4.5. Design for connection

Although users perceive their data as private, this does not mean that PI systems should impede their exchange with other individuals. Previous studies highlighted the opportunity to be social around data: Consolvo et al. (2006), for example, noted that when sharing activity level users feel a pressure to make their goal, as well as enjoy receiving recognition and encouragement from others; Toscos et al. (2006) suggested that exchanging data with friends may increase users' motivation; while

Bales and Griswold (2011) proposed to exploit the influence of the user's social networks to make change in her behaviour. More recently, research on trackers' real practices further emphasised that they are inclined to use sharing features to compare, compete and support motivation (Fritz et al. 2014); and when the activity tracked is framed in a cooperative game it also may enhance users' performances (Chen and Pu 2014; Rooksby et al. 2015). Moreover, Gorm and Shklovski (2016) noted that health-related personal data, such as steps, may be a socially negotiated quantity in the context of step-counting campaigns addressed to promote healthy behaviour.

Our findings point to the importance of the social aspects in PI. However, they also highlight that sharing data can be a means for establishing a connection with others. Our participants prefer to share their data with significant others or within strict circles of users. This attitude also explains their reluctance in exposing their data on general Social Network Sites: a hesitancy that has also been observed in experienced users, who do not commonly publish their data on Twitter or Facebook (Fritz et al. 2014; Rooksby et al. 2014), confirming a general aversion to expose data concerning wellbeing on general Social Network Sites (Newman et al. 2011). These channels may not be the best venues for regularly publishing PI data, also because their audience may not be interested in receiving updates about information tracked and providing support (Epstein et al. 2015). On the same vein, for our participants sharing data is not an act of publicity, but mainly a way for exchanging something private. Sharing in PI systems, then, should be thought of as a feature that builds on top of already established close interpersonal relationships in the user's 'real' life, or as a means for helping her better understand herself through the experiences of similar users. This means that PI systems should provide private communication means for sending personal information to specific targeted individuals, or be designed for small private circles where data can be freely shared among their members. In the domain of location data, an example of how to achieve this mutual exchange is represented by CoupleVIBE (Bales, Li, and Griwsold 2011) a mobile application that automatically pushes a user's location information to her partner's mobile phone via vibrotactile cues. Users should be allowed to know which kinds of information they are exposing; select the level of their exposure and choose the specific individuals that have access to their data. As U24 stressed, 'I want to decide which kind of data and to whom. This is essential!'. Even if other research found that concerns about the personal nature of their data are not common among experienced trackers (Fritz et al. 2014), our participants, both experienced and naïve, valued their private nature: as data exchanging could be also perceived as a sort of a 'personal gift', or a private memory to share, users should be assured that their exchange will not jeopardise their privacy.

6. Limitations

We list here the main limitations of our work, in relation to the diary study, the system and the final evaluation.

6.1. Diary study

The small sample size of the diary study limits the generalisability of its findings. Moreover, we did not control the sample, apart from participants' experience with and interest in PI. All the participants were well educated with a disproportionate number of students (six, counting both undergraduate and graduate) and individuals with a high level of instruction, having a Ph.D. (3). Participants were also relatively affluent, and very young individuals (<19) or older people (>50) were not represented. So, we cannot claim that this sample may represent any general population. Rather we suggest that these participants, in their unique ways, provide good cases of potential users of PI that could use these technologies in the future, given their interest and curiosity in discovering something about themselves through technology.

6.2. Evaluation study

The summative study suffered from limitations in ways similar to the diary study. We chose to conduct a qualitative research because we were primarily interested in assessing whether the requirements embedded in our system could help users find value in their data: we wanted to discover what they thought, and what kind of interpretation processes they enacted in exploring their data. To better highlight the study limitations, we will now discuss threats to the validity of the research, following Wholin et al. (2012).

There are four levels of validity threats to consider. *Internal validity* is primarily focused on the relationships between causes and effects studied in the research. *External validity* relates to the generalisation of the results from the context in which the study is conducted to a wider context. *Construct validity* concerns which components in a complex treatment really caused the effect. Finally, *conclusion validity* is related to the size of data sets, leading to statistical significance and to the ability to draw conclusions about relationship between treatment and outcome (Wholin et al. 2012).

With reference to *internal validity* the study primarily aimed at assessing the usefulness of the five user requirements found in the diary study, embedded within a PI system. Participants of two different groups emphasised that the key aspects of the system implementing the aforementioned requirements help them find value in their own data. The study, therefore, seemingly provides good internal validity about this point. The two groups showed different needs and goals, related to their level of expertise in self-tracking: however, they almost unanimously reported improvements in how they explored and valued personal data by using our system.

Instead, we cannot claim the same internal validity about the system's efficacy in producing a change in behaviour. Evaluating the system's capability of changing behaviour was not our primary goal. Nevertheless, we explored whether the system could affect users' intention to change. To truly demonstrate that a technology achieved its goal in modifying behaviour requires large studies with control groups that are typically not feasible for HCI technologies not fully deployed and commercialised (Klasnja, Consolvo, and Pratt 2011). Therefore, we preferred to tailor our evaluation to the assessment of a specific intervention strategy, the consciousness raising (i.e. the capability of eliciting the self-awareness about a behaviour that needs to be changed) (Prochaska and Velicer 1997), which usually provides benefits in the early stage of behaviour change process, as stated by the Trans Theoretical Model of behaviour change (Prochaska and Velicer 1997). Despite the useful findings we collected about this point, we have to clarify that our study lacked quantitative measures. We also did not set any comparison/control condition, and we did not tie the evaluation to specific health goals or clinical outcomes. These elements weaken our claims about the system's efficacy and capability of supporting behaviour change processes, even with regard to consciousness raising. However, as Klasnja, Consolvo, and Pratt (2011) noted, we think that a qualitative investigation that focuses on people's experiences with the technology could help researchers understand why and how the system is working. A larger and longer efficacy study could be undertaken in the future taking into account the findings that emerged during this research.

As regards the study's *external validity*, we better balanced the quota of students and professionals, as well as of the individuals having a Ph.D., with respect to the diary study. We also widened the age range (between 18 and 58) and increased the professions included in the sample. This might improve the study results' generalisability. However, almost all the participants (except the three workmen) were well educated and we did not include very young or very old people. They were also all relatively affluent. Moreover, almost all the experienced trackers involved in the study had a behaviour change goal. Although most of the experienced users do have a behaviour change goal (Choe et al. 2014; Fritz et al. 2014; Li, Dey, and Forlizzi 2010, 2011), our participants could not be representative of other experienced trackers with different motivations to track (Epstein et al. 2015; Rooksby et al. 2014).

The major threat regarding the *conclusion validity* is the quality of the data collected during the study. Participants did not integrate the system into their everyday routines to achieve their situated goals. Thus, there is a risk that the collected data are biased due to the study setting. However, during the think-aloud sessions, almost all the participants related the use of the system to real-life situations, imagining how it could support them in achieving their objectives, and comparing it with other self-tracking devices. This might suggest that the system could be easily integrated into their daily habits and that the feedback provided by the participants would not drastically change in a real context of use.

Finally, with respect to the *construct validity*, the system's features embedding the five user requirements seemed to increase the value participants found in their data. They spontaneously appreciated the system's functionalities implementing the user requirements pointing to how they improved their 'data experience'. However, with reference to the effectiveness of the system in supporting behaviour change processes we did not define appropriate measures to identify the system's aspects that may lead to a change in behaviour.

6.3. System's limitations

Although the data gathering was carried out on the field and the collected data were real, the system and the related visualisations were assessed in the lab. This was due to the technical limitations of our system, which was deployed on technologies that could not be easily placed in participants' homes (due to their costs and the space they take). Even if participants did not see this limitation as problematic in engaging with the visualisations, a continuous interaction with our system in a home environment would enable participant interest in continuously monitoring their changing trends.

PI systems/visualisations discussed in Section 2 satisfy only some of the requirements found in the dairy study, and often merely in a partial way, whereas our solution, as we have seen in Section 5, is capable to fully address all of them. However, it has still some limitations in comparison with previous work. First, differently from Consolvo et al.'s (2008a), Bauer et al.'s (2012), and

Bentley et al.'s (2013) systems, ours is not designed to provide just-in-time reminders or recommendations about specific behaviours: this is due to the fact that it was not primarily designed to change behaviour, rather provide users with a tool for reflection. Results from our study related to behaviour change, nonetheless, suggest that a further support in this direction would be needed. Second, our system does not fully consider dynamic querying and filtering: participants can only navigate in time but are not allowed, as in Baur et al.'s (2010) work, to select specific value ranges to filter data. This design decision has to be retraced to the need of keeping the interface as simple and engaging as possible: however, since some participants highlighted the need of having more fine-grained tools for analysing data, this should be considered in future development of the system. Third, the proposed system has not been thought of as an open web services infrastructure that would allow researchers to customise the interface and extend the tools available, as Salud! (Medynskiy and Mynatt 2010). Finally, with reference to context, our system does not provide information about people met, as Li, Dey, and Forlizzi's (2012) system does, or current calendar activities, as in Huang et al.'s work (2016). Such information could be considered for future work.

7. Conclusion and future work

Findings from our user study highlighted that our system makes PI data closer to PI users, enhances their value and improves self-reflection. The study also pointed out some acceptability and usability problems to fix. Especially, the 80″ screen raised questions about its integration in the users' everyday environments. Although we initially thought that reflecting the user in a natural-size body image would have increased her projection into their data, we discovered that this requirement was not necessary. Users individuated in their reflected body movements the main factor for identifying with their data. We will then experiment the display of the visualisation environments on 'normal-size' screens, such as TVs or desktop monitors.

Another point to be explored in the future is related to the analysis of the collected data. Our system gathers different types of time series data (i.e. data collected sequentially over time), such as the users' heartbeat and body temperature. Often in data analysis, we assume that observations are independent, but with time series data, this assumption may be false and time series analysis may account for this temporal correlation. However, currently the Data Analysis module of our system does not exploit methods for time series analysis. As future work, it would be interesting to investigate time series in our system. Time series analysis on PI data could reveal patterns such as random, periods, cycles, trends and unusual observations. The identification of trends and seasonal variations could enlighten hidden aspects of users' life, making them aware of how their parameters and behaviours may vary over time: a trend is an evolutionary movement, either upward or downward, in the data values and it may be dynamic and of short duration, or long term; while seasonality is the component of time series behaviour that regularly repeats, for example, each month. For a time series without any systematic change in its mean and variance and without periodic variations (stationary), or from which trends and seasonality have been removed ('stationarised'), we could investigate its correlational structure, exploring the extent to which successive terms in the series are correlated. This could be used to make predictions about the future value of the user's PI data.

To summarise, in this paper, we made three primary contributions. First, we presented a PI system designed around five user requirements specifically addressed to users with no previous experience in self-tracking. Second, we validated such requirements through a user study showing how they might ease and enhance the exploration of personal data for naïve users, as well as expert trackers. Third, we further individuated three general design considerations for PI that could orient the implementation of our future work and, by and large, the design of future PI systems.

Notes

1. One of them was also a software developer.
2. One of them was an e-commerce operator.
3. For some operations related to data acquisition/analysis/abstraction, we used the NITE library as well as the OpenNI library (https://github.com/OpenNI/OpenNI), wrapped in the simple-openni library (https://code.google.com/archive/p/simple-openni/).

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Ancker, J. S., and D. Kaufman. 2007. "Rethinking Health Numeracy: A Multidisciplinary Literature Review." *Journal of the American Medical Informatics Association* 14: 713–721.

Bales, E., and W. Griswold. 2011. "Interpersonal Informatics: Making Social Influence Visible." In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*, 2227–2232. New York, NY: ACM.

Bales, E., K. A. Li, and W. Griwsold. 2011. “CoupleVIBE: Mobile Implicit Communication to Improve Awareness for (Long-distance) Couples.” In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW ‘11)*, 65–74. New York, NY: ACM.

Bauer, J. S., S. Consolvo, B. Greenstein, J. Schooler, E. Wu, N. F. Watson, and J. Kientz. 2012. “ShutEye: Encouraging Awareness of Healthy Sleep Recommendations with a Mobile, Peripheral Display.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1401–1410. New York, NY: ACM.

Baur, D., F. Seiffert, M. Sedlmair, and S. Boring. 2010. “The Streams of Our Lives: Visualizing Listening Histories in Context.” *IEEE Transactions on Visualization and Computer Graphics* 16 (6): 1119–1128.

Bentley, F., T. Tollmar, P. Stephenson, L. Levy, B. Jones, S. Robertson, E. Price, R. Catrambone, and J. 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) (Article 30, 27 pages).

Bergman, L., B. E. Rogowitz, and L. A. Treinish. 1995. “A Rule-based Tool for Assisting Colormap Selection.” In *Proceedings of IEEE Visualization*, Atlanta, GA, 118–125.

Bertini, E., S. Gabrielli, and S. Kimani. 2006. “Appropriating and Assessing Heuristics for Mobile Computing.” In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI ‘06)*, 119–126. New York, NY: ACM.

Birk, M. V., C. Atkins, J. T. Bowey, and R. L. Mandryk. 2016. “Fostering Intrinsic Motivation Through Avatar Identification in Digital Games.” In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI ‘16)*, 2982–2995. New York, NY: ACM.

Borland, D., and R. M. Taylor II. 2007. “Rainbow Color Map (Still) Considered Harmful.” *IEEE Computer Graphics and Applications* 27 (2): 14–17.

Bowen, G. A. 2008. “Naturalistic Inquiry and the Saturation Concept: A Research Note.” *Qualitative Research* 8 (1): 137–152.

Braun, V., and V. Clarke. 2006. “Using Thematic Analysis in Psychology.” *Qualitative Research in Psychology* 3: 77–101.

Brewer, C. 1994. “Color Use Guidelines for Mapping and Visualization.” In *Visualization in Modern Cartography*, edited by A. M. MacEachren and D. R. F. Taylor, 123–147. Tarrytown, NY: Elsevier Science.

Brewer, C. 1999. “Color Use Guidelines for Data Representation.” In *Proceedings of the Section on Statistical Graphics*, 55–60. Alexandria, VA: American Statistical Association.

Carter, S., and J. Mankoff. 2005. “When Participants do the Capturing: The Role of Media in Diary Studies.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘05)*, 899–908. New York, NY: ACM.

Cena, F., S. Likavec, and A. Rapp. 2015. “Quantified Self and Modeling of Human Cognition.” In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC’15 Adjunct)*, 1021–1026. New York: ACM.

Cena, F., I. Lombardi, A. Rapp, and F. Sarzotti. 2014. “Self-monitoring of Emotions: A Novel Personal Informatics Solution for an Enhanced Self-reporting.” In *Workshop Proceedings of the 22nd Conference on User Modeling, Adaptation and Personalization UMAP ‘14. CEUR Workshop Proceedings*, Vol. 1181.

Chen, Y., and P. Pu. 2014. “HealthyTogether: Exploring Social Incentives for Mobile Fitness Applications.” In *Proceedings of the Second International Symposium of Chinese CHI (Chinese CHI ‘14)*, 25–34. New York, NY: ACM.

Cheng, W. C., L. Golubchik, and D. G. Kay. 2004. “Total Recall: are Privacy Changes Inevitable?” In *Proceedings of the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE ‘04)*, 86–92. New York, NY: ACM.

Choe, E. K., N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz. 2014. “Understanding Quantified-selfers’ Practices in Collecting and Exploring Personal Data.” In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems (CHI ‘14)*, 1143–1152. New York, NY: ACM.

Church, K., and B. Smyth. 2009. “Understanding the Intent Behind Mobile Information Needs.” In *Proceedings of the 14th International Conference on Intelligent User Interfaces (IUI ‘09)*, 247–256. New York, NY: ACM.

Clawson, J., J. A. Pater, A. D. Miller, E. D. Mynatt, and L. Mamykina. 2015. “No Longer Wearing: Investigating the Abandonment of Personal Health-Tracking Technologies on Craigslist.” In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp ‘15)*, 647–658. New York: ACM.

Consolvo, S., K. Everitt, I. Smith, and J. Landay. 2006. “Design Requirements for Technologies That Encourage Physical Activity.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘06)*, 457–466. New York, NY: ACM.

Consolvo, S., P. Klasnja, D. W. McDonald, D. Avrahami, J. Froelich, L. LeGrand, R. Libby, K. Mosher, and J. A. Landay. 2008a. “Flowers or a Robot Army? Encouraging Awareness & Activity with Personal, Mobile Displays.” In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp ‘08)*, 54–63. New York, NY: ACM.

Consolvo, S., D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, et al. 2008b. “Activity Sensing in the Wild: a Field Trial of Ubifit Garden.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘08)*, 1797–1806. New York, NY: ACM.

Czerwinski, M., E. Horvitz, and S. Wilhite. 2004. “A Diary Study of Task Switching and Interruptions.” In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ‘04)*, 175–182. New York, NY: ACM.

Elliot, A. J., R. G. Miltenberger, J. Kaster-Bundgaard, and V. Lumley. 1996. “A National Survey of Assessment and Therapy Techniques Used by Behavior Therapists.” *Cognitive and Behavioral Practice* 3: 107–125.

Elsden, C., D. S. Kirk, and A. C. Durrant. 2016. “A Quantified Past: Toward Design for Remembering with Personal Informatics.” *Human–Computer Interaction* 31 (6): 518–557.

Epstein, D., F. Cordeiro, E. Bales, J. Fogarty, and S. Munson. 2014. “Taming Data Complexity in Lifelogs: Exploring

Visual Cuts of Personal Informatics Data." In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*, 667–676. New York, NY: ACM.

Epstein, D. A., A. Ping, J. Fogarty, and S. A. Munson. 2015. "A Lived Informatics Model of Personal Informatics." In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*, 731–742. New York, NY: ACM.

Etter, J. F., and S. Sutton. 2002. "Assessing 'Stage of Change' in Current and Former Smokers." *Addiction* 97: 1171–1182.

Feltron, N. 2014. "Feltron's Annual Report." Accessed November 4, 2017. http://feltron.com/

Foster, S. L., C. Laverty-Finch, D. P. Gizzo, and J. Osantowski. 1999. "Practical Issues in Self-observation." *Psychological Assessment* 11: 426–438.

Fritz, T., E. M. Huang, G. C. Murphy, and T. Zimmermann. 2014. "Persuasive Technology in the Real World: A Study of Long-term Use of Activity Sensing Devices for Fitness." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, 487–496. New York, NY: ACM.

Galesic, M., and R. Garcia-Retamero. 2011. "Graph Literacy: A Crosscultural Comparison." *Medical Decision Making* 31: 444–457.

Gemmell, J., G. Bell, and R. Lueder. 2006. "MyLifeBits: A Personal Database for Everything." *Communications of the ACM* 49 (1): 88–95.

Gobo, G. 2004. "Sampling, Representativeness and Generalizability." In *Qualitative Research Practice*, edited by C. Seale, G. Gobo, J. Gubrium, and D. Silverman, 435–456. London: Sage.

Gorm, N., and I. Shklovski. 2016. "Steps, Choices and Moral Accounting: Observations From a Step-counting Campaign in the Workplace." In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*, 148–159. New York, NY: ACM.

Gouveia, R., E. Karapanos, and M. Hassenzahl. 2015. "How Do We Engage with Activity Trackers?: A Longitudinal Study of Habito." In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 1305–1316. New York, NY: ACM.

Gouveia, R., F. Pereira, E. Karapanos, S. A. Munson, and M. Hassenzahl. 2016. "Exploring the Design Space of Glanceable Feedback for Physical Activity Trackers." In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*, 144–155. New York, NY: ACM.

Grinter, R., and M. Eldridge. 2003. "Wan2tlk?: Everyday Text Messaging." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*, 441–448. New York, NY: ACM.

Hekler, E. B., P. Klasnja, J. E. Froehlich, and M. P. Buman. 2013. "Mind the Theoretical gap: Interpreting, Using, and Developing Behavioral Theory in HCI Research." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, 3307–3316. New York: ACM.

Hilviu, D., and A. Rapp. 2015. "Narrating the Quantified Self." In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*, 1051–1056. New York: ACM.

Ho, H.-N., G. H. Van Doorn, T. Kawabe, J. Watanabe, and C. Spence. 2014. "Colour–Temperature Correspondences: When Reactions to Thermal Stimuli are Influenced by Colour." *PLoS ONE* 9 (3): e91854.

Huang, D., M. Tory, B. A. Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury. 2015. "Personal Visualization and Personal Visual Analytics." *IEEE Transactions on Visualization and Computer Graphics* 21 (3): 420–433.

Huang, D., M. Tory, and L. Bartram. 2016. "A Field Study of On-calendar Visualizations." In *Proceedings of the 42nd Graphics Interface Conference (GI '16)*, Canadian Human-Computer Communications Society, Waterloo, Ontario, 13–20.

Hunter, L. 2014. "Are Wearables Over? Fast Company Inc." Accessed March 1, 2016. http://www.fastcompany.com/3028879/most-innovative-companies/are-wearables-over.

IDTechEx. 2014. "Wearable Technology 2014-2024: Technologies, Markets, Forecasts E-Textiles, Wearable Electronics, Medical Diagnostics, Smart Glasses, Smart Wristbands and More." Accessed March 1, 2016. https://www.prnewswire.com/news-releases/wearable-technology-2014-2024-technologies-markets-forecasts-254355281.html.

Karkar, R., J. Fogarty, J. A. Kientz, S. A. Munson, R. Vilardaga, and J. Zia. 2015. "Opportunities and Challenges for Self-experimentation in Self-tracking." In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*, 991–996. New York, NY: ACM.

Kawamoto, K., K. Hiroyuki, and T. Seiki. 2013. "Actigraphic Detection of REM Sleep Based on Respiratory Rate Estimation." *Journal of Medical and Bioengineering* 2 (1): 20–25.

Kawamoto, K., T. Tanaka, and H. Kuriyama. 2014. "Your Activity Tracker Knows When you Quit Smoking." In *Proceedings of the 2014 ACM International Symposium on Wearable Computers (ISWC '14)*, 107–110. New York, NY: ACM.

Kay, M., E. K. Choe, J. Shepherd, B. Greenstein, N. Watson, S. Consolvo, and J. A. Kientz. 2012. "Lullaby: A Capture & Access System for Understanding the Sleep Environment." In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (Ubicomp '12)*, 226–234. New York, NY: ACM.

Kim, S., E. Paulos, and J. Mankoff. 2013. "inAir: A Longitudinal Study of Indoor air Quality Measurements and Visualizations." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2745–2754. New York, NY: ACM.

Klasnja, P., S. Consolvo, and W. Pratt. 2011. "How to Evaluate Technologies for Health Behavior Change in HCI Research." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 3063–3072. New York, NY: ACM.

Korotitisch, W. J., and R. O. Nelson-Gray. 1999. "An Overview of Self-monitoring Research in Assessment and Treatment." *Psychological Assessment* 11: 415–425.

Lazar, A., C. Koehler, J. Tanenbaum, and D. H. Nguyen. 2015. "Why We Use and Abandon Smart Devices." In *Proceedings*

of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15), 635–646. New York, NY: ACM.
Ledger, D., and D. McCaffrey. 2014. "Inside Wearables: How the Science of Human Behavior Change Offers the Secret to Long-term Engagement." *Endeavour Partners LLC* 93 (2014): 36–45.
Lee, M. K., J. Kim, J. Forlizzi, and S. Kiesler. 2015. "Personalization Revisited: A Reflective Approach Helps People Better Personalize Health Services and Motivates Them to Increase Physical Activity." In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*, 743–754. New York, NY: ACM.
Li, I., A. K. Dey, and J. Forlizzi. 2010. "A Stage-based Model of Personal Informatics Systems." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*, 557–566. New York, NY: ACM.
Li, I., A. K. Dey, and J. Forlizzi. 2011. "Understanding My Data, Myself: Supporting Self-reflection with Ubicomp Technologies." In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*, 405–414. New York, NY: ACM.
Li, I., A. K. Dey, and J. Forlizzi. 2012. "Using Context to Reveal Factors That Affect Physical Activity." *ACM Transactions on Computer-Human Interaction* 19 (1) (Article 7, 21 pages).
Liu, W., B. Ploderer, and T. Hoang. 2015. "In Bed with Technology: Challenges and Opportunities for Sleep Tracking." In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '15)*, 142–151. New York, NY: ACM.
Lu, H., D. Frauendorfer, M. Rabbi, M. Schmid Mast, G. T. Chittaranjan, A. T. Campbell, D. Gatica-Perez, and T. Choudhury. 2012. "StressSense: Detecting Stress in Unconstrained Acoustic Environments Using Smartphones." In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12)*, 351–360. New York, NY: ACM.
Lupton, D. 2016. "You are Your Data: Self-tracking Practices and Concepts of Data." In *Lifelogging: Digital Self-Tracking and Lifelogging – Between Disruptive Technology and Cultural Transformation*, edited by S. Selke, 61–70. Wiesbaden: Springer Fachmedien.
MacQueen, K. M., E. McLellan-Lemal, K. Bartholow, and B. Milstein. 2008. "Team-based Codebook Development: Structure, Process, and Agreement." In *Handbook for Team-based Qualitative Research*, edited by G. Guest and K. M. MacQueen, 119–135. Lanham, MD: AltaMira Press.
Maddox, T. 2014. "Wearables Have A Dirty Little Secret: 50% of Users Lose Interest." Tech Republic Inc. Accessed March 1, 2016. http://www.techrepublic.com/article/wearables-have-a-dirty-little-secret-most-people-lose-interest/.
Mann, S. 2004. "Continuous Lifelong Capture of Personal Experiences with EyeTap." In *Proceedings of the 1st ACM Workshop on Continuous Archival and Retrieval of Personal Experiences (CARPE '04)*, 1–21. New York, NY: ACM.
Marcengo, A., and A. Rapp. 2014. "Visualization of Human Behavior Data: The Quantified Self." In *Innovative Approaches of Data Visualization and Visual Analytics*, edited by L. H. Huang and W. Huang, 236–265. Hershey, PA: IGI Global.
Marshall, M. N. 1996. "Sampling for Qualitative Research." *Family Practice* 13 (6): 522–526.
Matthews, M., and G. Doherty. 2011. "In the Mood: Engaging Teenagers in Psychotherapy Using Mobile Phones." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*, 2947–2956. New York, NY: ACM.
Medynskiy, Y., and E. Mynatt. 2010. "Salud!: An Open Infrastructure for Developing and Deploying Health Self-management Applications." *Proceedings of Pervasive Health* 10: 1–8.
Miltenberger, R. G. 2007. *Behavior Modification: Principles and Procedures*. 4th ed. Belmont, MA: Wadsworth.
Mori, M., K. F. MacDorman, and N. Kageki. 2012. "The Uncanny Valley [From the Field]." *Robotics & Automation Magazine, IEEE* 19 (2): 98–100.
Newman, M., D. Lauterbach, S. Munson, P. Resnick, and M. Morris. 2011. ""It's Not That I Don't Have Problems, I'm Just Not Putting Them on Facebook": Challenges and Opportunities in Using Online Social Networks for Health." In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*, 341–350. New York, NY: ACM.
Nielsen, J. 1994. "Heuristic Evaluation." In *Usability Inspection Methods*, edited by J. Nielsen and R. L. Mack, 25–62. New York, NY: Wiley.
Pham, C., D. Jackson, J. Schoening, T. Bartindale, T. Ploetz, and P. Olivier. 2013. "FoodBoard: Surface Contact Imaging for Food Recognition." In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*, 749–752. New York, NY: ACM.
Pousman, Z., J. Stasko, and M. Mateas. 2007. "Casual Information Visualization: Depictions of Data in Everyday Life." *Proceedings of IEEE InfoVis 2007*. Sacramento, CA.
Prekop, P., and M. Burnett. 2003. "Activities, Context and Ubiquitous Computing." *Computer Communications* 26 (11): 1168–1176.
Prochaska, J. O., and W. F. Velicer. 1997. "The Transtheoretical Model of Health Behavior Change." *American Journal of Health Promotion* 12 (1): 38–48.
Rapp, A., and F. Cena. 2014. "Self-monitoring and Technology: Challenges and Open Issues in Personal Informatics." In *Proceedings of HCI International. Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice*, Vol. 8516, 613–622. Lecture Notes in Computer Science. Cham: Springer.
Rapp, A., and F. Cena. 2015. "Affordances for Self-tracking Wearable Devices." In *Proceedings of International Symposium on Wearable Computers, ISWC 2015*, 141–142. New York: ACM.
Rapp, A., and F. Cena. 2016. "Personal Informatics for Everyday Life: How Users without Prior Self-Tracking Experience Engage with Personal Data." *International Journal of Human-Computer Studies* 94: 1–17.
Rapp, A., F. Cena, J. Kay, B. Kummerfeld, F. Hopfgartner, T. Plumbaum, and J. E. Larsen. 2015. "New Frontiers of Quantified Self: Finding New Ways for Engaging Users in Collecting and Using Personal Data." In *Adjunct Proceedings of the 2015 ACM International Joint*

Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct), 969–972. New York: ACM. doi:10.1145/2800835.2807947. ISBN: 978-1-4503-3575-1

Rapp, A., F. Cena, J. Kay, B. Kummerfeld, F. Hopfgartner, T. Plumbaum, J. E. Larsen, D. A. Epstein, and R. Gouveia. 2016. "New Frontiers of Quantified Self 2: Going Beyond Numbers." In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*, 506–509. New York: ACM.

Rapp, A., F. Cena, J. Kay, B. Kummerfeld, F. Hopfgartner, T. Plumbaum, J. E. Larsen, D. A. Epstein, and R. Gouveia. 2017. "New Frontiers of Quantified Self 3: Exploring Understudied Categories of Users." In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17)*, 861–864. New York: ACM.

Rapp, A., and M. Tirassa. 2017. "Know Thyself: A Theory of the Self for Personal Informatics." *Human-Computer Interaction* 32 (5–6): 335–380.

Reas, C., and B. Fry. 2007. *Processing, a Programming Handbook for Visual Designers and Artists*. Cambridge, MA: The MIT Press.

Rooksby, J., M. Rost, A. Morrison, and M. Chalmers. 2014. "Personal Tracking as Lived Informatics." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, 1163–1172. New York, NY: ACM.

Rooksby, J., M. Rost, A. Morrison, and M. Chalmers. 2015. "Pass the Ball: Enforced Turn-taking in Activity Tracking." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, 2417–2426. New York, NY: ACM.

Schmidt, B., S. Benchea, R. Eichin, and C. Meurisch. 2015. "Fitness Tracker or Digital Personal Coach: How to Personalize Training." In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct)*, 1063–1067. New York, NY: ACM.

Sohn, T., K. A. Li, W. G. Griswold, and J. D. Hollan. 2008. "A Diary Study of Mobile Information Needs." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, 433–442. New York, NY: ACM.

Strauss, A. L., and J. M. Corbin. 1990. "Grounded Theory Research: Procedures, Canons, and Evaluative Criteria." *Qualitative Sociology* 13 (1): 3–21.

Sumida, M., T. Mizumoto, and K. Yasumoto. 2013. "Estimating Heart Rate Variation During Walking with Smartphone." In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*, 245–254. New York, NY: ACM.

Tene, O., and J. Polonetsky. 2013. "A Theory of Creepy: Technology, Privacy and Shifting Social Norms." *Yale Journal of Law & Technology* 16: 59–134.

Toscos, T., A. Faber, S. An, and M. P. Gandhi. 2006. "Chick Clique: Persuasive Technology to Motivate Teenage Girls to Exercise." In *CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06)*, 1873–1878. New York, NY: ACM.

Tulving, E. 2002. "Episodic Memory: From Mind to Brain." *Annual Review of Psychology* 53: 1–25.

Warshaw, J., T. Matthews, S. Whittaker, C. Kau, M. Bengualid, and B. A. Smith. 2015. "Can an Algorithm Know the "Real You"?: Understanding People's Reactions to Hyper-personal Analytics Systems." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*, 797–806. New York: ACM.

Wholin, C., P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. 2012. *Experimentation in Software Engineering*. Berlin, Heidelberg: Springer.

Whooley, M., B. Ploderer, and K. Gray. 2014. "On the Integration of Self-tracking Data Amongst Quantified Self Members." In *Proceedings of the International BCS Human Computer Interaction Conference (BCS-HCI '14)*, Southport, 151–160.

Appendix

A.1. Diary study

Screening criteria. The inclusion criteria included participants who (i) were curious or interested in self-monitoring a specific parameter/behaviour through technology.

The exclusion criteria included participants who (i) were currently using one or more PI tools; (ii) had used PI tools in the past; (iii) had a specific need to self-track, such as a specific aspect of their life that they wanted to change, or a chronic disease.

Diary prompts. Personal considerations on the daily use of the provided instruments; Issues faced in using the provided instruments; Insights gained and discoveries found during the exploration of the collected data; Reflections elicited by the exploration of the data; Consequences of self-monitoring on daily habits; Breakdowns and unforeseen episodes.

Semi-structured interview questions. Could you describe your overall experience with the provided instruments? Could you list the positive and the negative aspects of such experience? Would you use these tools in your everyday life, how and why? Do you have any reflection on how these tools could be improved in the future?

A.2. Evaluation study

Screening criteria with regard to experience and interest in PI. The inclusion criteria for naïve users included participants who (i) were curious or interested in self-monitoring a particular parameter/behaviour through technology.

The exclusion criteria included participants who (i) were currently using one or more PI tools; (ii) had used PI tools in the past.

The inclusion criteria for experienced users included participants who (i) had been using one or more self-tracking tools for more than three months.

The exclusion criteria included participants who (i) had stopped to self-track before the beginning of the study.

Assessment of stages of change with regard to health behaviour

(i) whether they considered a particular behaviour related to their health (i.e. eating habits, sleep, physical activity levels, stress) problematic and thus requiring a change (if not, precontemplation stage);
(ii) whether they were seriously thinking to change a particular behaviour within the next the next 6 months or within the next 30 days and whether they had already attempted to change that behaviour in the past (contemplation stage and preparation stage);
(iii) whether they were currently trying to change a specific behaviour (within the last six months) (action stage);
(iv) whether they were trying to maintain a certain behaviour that they had changed in the past (more than six months ago) (maintenance stage).

Semi-structured interview questions. Could you describe your overall experience with the system? Could you detail difficulties, positive elements and personal reflections about the system? What do you think about the evocative visualisation? and about the synthetic visualisation? What kind of further considerations could you make looking at your data within our system? Did seeing these data change something about what you thought of yourself? Do you have any idea on how this system could be improved in the future?