Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data

Published version:

DOI:10.1016/j.ijhcs.2016.05.006

Personal Informatics for Everyday Life: How Users without Prior Self-Tracking Experience Engage with Personal Data

Amon Rapp[1], Federica Cena[2]

University of Torino – Computer Science Department, C.so Svizzera, 185 Torino, 10149 Italy

[1]amon.rapp@gmail.com; [2]cena@di.unito.it

ABSTRACT

The spreading of devices and applications allowing people to collect personal information opens new opportunities for Personal Informatics. Although many of these tools are already effectively used by motivated people to gain self-knowledge and produce change in their behaviors, there is a great number of users that are potentially interested in Personal Informatics but do not know of its potentialities and criticalities. In order to investigate how users perceive and use self-tracking tools in everyday life, we conducted a diary study, requiring fourteen participants with no previous experience in Personal Informatics to use a variety of trackers. We discovered that they use and perceive these technologies differently from the ones experienced in self-tracking. Participants considered the act of collecting personal information burdensome, with no beneficial reward. We also uncovered a series of problems that they experienced while tracking, managing, visualizing, and using their data. Among them we found that the lack of suggestions on using data and the excess of abstract visualization in the apps prevented users to gain useful insights. As a result, their interest in self-tracking soon faded, despite their initial curiosity in exploring and "playing" with their data. Starting from the findings of this study, we identified seven design strategies to better Personal Informatics tools, supported by literature and examples that draw from different research fields, from tangible interfaces, to virtual environments and video games. These strategies are primarily addressed to satisfy the inexperienced users' needs, but their applicability can be reasonably extended to all the individuals curious and interested in Personal Informatics.

Keywords

Personal informatics, quantified self, diary study, design strategies.

Contact author: Amon (first name) Rapp (last name)

University of Torino – Computer Science Department

C.so Svizzera, 185 Torino, 10149 Italy

amon.rapp@gmail.com; rapp@di.unito.it

Mobile: +39 3462142386

1. INTRODUCTION

The new possibilities offered by technological advances in sensors and portable devices open new opportunities for Personal Informatics (PI) systems, "those that help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge" (Li et al., 2010). Nowadays, a large variety of data can be gathered by means of ubiquitous and wearable technologies: from internal states (as mood or glucose level in the blood) to performance values (as pace or kilometers run), from habits (as food, sleep) to actions (as visited places). PI tools allow users to *self-monitor* their behaviors in a variety of contexts simplifying data collection, management and visualization.

Gathering personal data and reflecting on them has a long history. For medical and clinical purposes self-monitoring systems were used for years under the guidance of therapists or physicians (Elliot et. al, 1996). Nevertheless, other technologies for self-tracking, such as weight scales, shifted from the doctor's office to the home, from a form of specialist medical knowledge to a private habit and an everyday domestic discipline, since the start of the 20th century (Crawford et al., 2015); while other instruments, like pedometers, were used privately to meet fitness goals since their appearance on the commercial market in the middle of the last century (Tudor-Locke and Bassett, 2004).

Personal Informatics systems gained popularity in the last years first among researchers and technology enthusiasts, such as the members of the Quantified Self movement, which tracked their everyday behaviors with the aim of raising their self-awareness (Marcengo and Rapp, 2013). Nowadays, we are assisting to their diffusion at consumer level, widening the possibility of their widespread adoption and their integration in everyday practices. On the one side, a plethora of wearable devices is entering the market: some predictions indicate that wearable personal-tracking technologies will eclipse $70 Billion by 2024 (IdTechEx, 2014). On the other side, the pervasiveness of self-tracking in modern smartphones foreshadows an era where Personal Informatics will likely become ubiquitous, making personal data available with minimal burden, easing the process of self-monitoring (Klansja & Pratt, 2012).

As long as personal data will be easily available to people, new users will likely become interested in Personal Informatics. There is a huge user base made up of "potential users" that could find, in the near future, a motivation in tracking, managing and visualizing their personal data. This enlarged user base can be represented by healthy and health-conscious individuals, or unhealthy but who strive to be healthier, or consumers of organic products and likely to exercise in spare time, young, educated and technology savvy, sharing thus some of the characteristics that pertain to users that are currently showing interest in wearables (Salah et al., 2014). However, market predictions highlight also that these peculiarities will soon no longer describe the totality of self-trackers (Epstein et al., 2015). All those people that actually own a smartphone, are sufficiently open to technology to try a new device or application, and show curiosity or interest in understanding something about themselves could represent this new user base.

This nonetheless poses a double problem. First, these "new" potential users are unfamiliar with PI technologies and may have a misperception of their limits and potentialities, by not knowing exactly what kind of efforts they may require and benefits they may provide. It is probable that they are not inclined to invest a huge amount of time for managing and understanding their data. These potential users also likely do not have specific goals for tracking at the beginning of their experience with PI tools; however, they could find their own objectives during the tracking activity, provided that these instruments could show them the value of the collected data for their everyday life, by stimulating tacit desires or raising unexpressed needs.

Second, we believe that current PI tools are not yet designed with enough understanding of these users' needs, desires and problems that they may encounter. There exists a growing skepticism regarding these technologies and their ability to provide concrete benefits (Hammond, 2014; Hunter, 2014), or inspire and sustain individuals' engagement in everyday life (Herz, 2014). A recent study reports that one third of the Americans that purchased a self-monitoring tool abandoned it after 6 months of use (Ledger & McCaffrey, 2014).

For these reasons, it is essential to start to think about how we could design for making these technologies, and the data gathered by them, interesting also for those users that could adopt them in the near future. To understand whether these instruments can meet their expectations, or fail due to issues in their designs, we conducted a diary study, requesting users with no previous experience in PI to use different PI applications and devices during their daily practices. Up until now, some studies were conducted in order to discover how PI tools are used, but they were mainly focused on individuals already familiar with such instruments (e.g. Fritz at al. 2014, Li et al., 2010, Li et al., 2011).

With this work, instead, we want to provide the following contributions: i) to individuate barriers that people with no experience with PI may find in tracking, managing, visualizing, and using personal data; ii) to describe how they use and perceive current PI technologies, highlighting the differences with respect to experienced users; iii) to define design strategies to address some of the main relevant issues found.

The relevance of this work is twofold. First, understanding how potential users of PI use self-tracking tools can highlight the issues that they may encounter and reveal unsatisfied needs and desires on which to focus design efforts, in order to address this class of instruments to an enlarged user base. Second, a series of design suggestions will allow developers to create more useful tools that could engage new users in self-tracking, allowing them to discover the utility behind their personal data.

The paper is structured as follows. First, we provide an overview of the previous studies related to the usage of PI in Section 2. Then, in Section 3, we describe the methodology followed in our research and, in Section 4, its results. Section 5 discusses our findings with respect to similar works. Finally, Section 6 proposes a series of design strategies to address some of the identified issues. Section 7 concludes the paper.

2. RELATED WORK

Self-monitoring, the activity of observing and recording one's own behavior (i.e., actions, thoughts and emotions), is a well-known technique in cognitive and behavioral psychology (Foster et al., 1999, Korotitisch and Nelson-Gray, 1999). Conceived as a clinical assessment method for collecting data on behaviors that only the patient could observe and record (e.g. eating, smoking), self-monitoring has become a standalone intervention technique, because of its reactive effects. Reactivity refers to the phenomenon in which the process of recording behavior causes the behavior to change (Nelson and Hayes 1981): self-monitoring often changes behavior, and this change is typically in the desired direction (Miltenberger, 2007). Personal Informatics technologies enhance the self-monitoring process, allowing people to track their behaviors outside the clinical setting.

Personal Informatics was first defined by Li et al. (2010) as a class of applications that help people collect and reflect on personal information. Pioneering examples of PI applications can be found in life logging research (Cheng et al, 2004, Gemmell et al, 2006, Mann et al., 2004), "a form of pervasive computing consisting of a unified digital record of the totality of an individual's experiences" (Dodge & Kitchin, 2007). As years passed, other academic research aimed at designing systems that allow people to collect and visualize their personal information, for therapeutic and rehabilitation purposes, or for promoting behavior change towards healthier and more sustainable habits. For example, UbiFit Garden (Consolvo et al., 2008) uses wearable technologies and a personal, mobile display to encourage physical activity; UbiGreen (Froehlich et al., 2009) is a mobile application providing personal awareness about green transportation habits through iconic feedback; Mobile Mood Diary (Matthews and Doherty, 2011) is a mobile and online symptom tracking tool for adolescents with mental problems; Lullaby (Kay et al., 2012) is a system to track sleep that combines temperature, motion sensors, light, audio, photos and an off-the-shelf sleep sensor to help people understand their sleep behavior. During the same years, PI tools started to be used by groups of technology enthusiasts, such as the Quantified Selfers, with the aim of increasing their self-knowledge, or improving their everyday life through self-experimentation (Marcengo and Rapp, 2013).

As of today, the availability of many commercial applications and devices for self-tracking boosted the popularity of PI among a wider range of users. Healthcare and fitness seem now the most popular domains for Personal Informatics. On the one side, PI tools help users self-track the evolution of symptoms of specific diseases, like Asthmapolis that uses sensors to track asthma attacks, or the variation of particular physiological parameters, like SugarStats that monitors blood sugar levels. On the other side, wearables like Fitbit and Jawbone Up collect and analyze information related to users' physical activity and sleep behavior, while applications like MyFitnessPal, Loseit and Calorie Counter allow people to keep track of their food intakes by simply filling a diary. However, self-tracking instruments are rapidly spreading among a variety of domains, enabling the gathering of data related to mood and emotions (e.g. Expereal and T2 Mood Tracker), daily tasks (Daytum), movement and locations (Moves), dreams (Dreamboard), finance (Mint), and so on.

Some of these instruments track data automatically, by leveraging sensors and algorithms capable of inferring the target behavior or activity from the data collected. Others, instead, need to rely on the user's self-reporting, given the complexity of the parameter tracked. In Table 1 we try to summarize the variety of commercial PI tools in a taxonomy, based on the type of data that they track, with examples for each category. This taxonomy does not aim to be exhaustive, as there currently are myriads of instruments for self-tracking, and the types of information that they are capable of collecting are rapidly expanding. On the contrary, it is only meant to give a snapshot of what PI tools are.

Type of Data	**Self-reported**	**Automatically detected**	**Main Domain**
Psychological Parameters	Expereal (Mood)	Empatica (Stress / Arousal / Epilepsy)	Health /

Type of Data	Self-reported	Automatically detected	Main Domain
	T2 Mood Tracker (Mood)	Phyode W/Me (Emotions / Nervous System) InteraXon Muse (Mental Activity)	Wellness
Physiological Parameters	SugarStats (Blood sugar) MonthlyInfo (Menstruation)	Zio XT Patch (Heart Rate) Jawbone Up (Sleep Patterns) SleepBot (Sleep Patterns)	Health / Wellness
Symptoms	CatchMyPain (Pain)	Asthmapolis (Asthma attacks)	Health
Behaviors	MyFitnessPal (Food) Dreamboard (Dreams) Drinking diary (Alcohol)	Nike + (Run) Jawbone Up (Physical Activity) Lumo Fit (Posture)	Wellness / Fitness
Daily Tasks and Management	Daytum (Tasks) Kibotzer (Goals)	Mint (Finance)	Self-management
Movements / Locations	Mileage Log+ (Travels)	Moves (Movements)	Transportation

Table 1. An overview of different PI tools currently available on the market

These instruments are now available for anyone interested in tracking their own data and soon they will allow potentially everyone to collect personal information about herself. In this landscape, a question arises: are these technologies designed for this potentially enlarged user base?

Until now, research on the usage of available PI tools on the market was focused exclusively on how "experienced users", such as Quantified Selfers, use and perceive these kinds of technologies. Li et al. (2010), for example, carried out a survey with individuals who collect and reflect on PI. Based on its findings, they suggested a stage-based model of PI composed of a series of five stages through which trackers transition when using PI tools: *preparation*, where they start collecting personal information; ii) *collection*, when they gather data about themselves; iii) *integration*, where the information gathered are transformed for the user to reflect on; iv) *reflection*, when the user reflects on her information; v) *action*, when trackers choose how to behave thanks to their newfound self-understanding. Li et al. further identified barriers that participants experienced in each of these stages, highlighting, for example, how they found burdensome manually collecting data, and how they encountered difficulties in integrating data coming from multiple inputs.

Li et al. (2011) further investigated the usage of PI tools in another work showing that the current commercial tools do not have sufficient understanding of users' needs. These instruments, for example, did not help users explore their data holistically: therefore, some of their participants managed their data together by using paper graphs or by reviewing their data logs.

Fritz et al. (2014) studied “in the wild” the long-term usage of commercial wearable devices addressed to promote physical wellness, focusing on the behavior change processes and motivations. They described how self-trackers showed a strong attachment to their devices, and how these systems had immediate impacts on their activities. They further emphasized that long-term trackers were highly motivated by the numerical feedback provided by the devices and that the tools’ social features offered them a set of supplementary potential goals to improve their physical activity.

Choe et al. (2014), instead, interviewed an “extreme” group of users, the Quantified Selfers, classifying their motivations in tracking their own behaviors. They described how these individuals found different issues in tracking, managing and visualizing their own data, due to a variety of lacks in the tools’ features. However, they stressed how these users always searched for a solution that could meet their needs, by building, for example, their own tools for tracking, by formulating hypotheses on their own data to test, or by creating their own visualizations (such as line charts).

Finally, Rooksby et al. (2014) pointed out how users use differently diverse tracking technologies, identifying a variety of personal tracking styles: 1) *directive*, to achieve a goal; 2) *documentary*, to record activities 3) *diagnostic*, to link different parameters together; 4) *collecting rewards*, to gather incentives; and 5) *fetishised*, for a pure interest in data or technology. The authors also introduced the expression ‘lived informatics’ to outline the people’s real practices in tracking information, meaning with it that “people are using information and finding its meaning in their day-to-day lives” (Rooksby et al., 2014).

However, all these studies involved participants with a strong previous experience with Personal Informatics. Li et al. (2010), for example, recruited participants from blogs and forums dedicated to QS, information visualization and PI, stating that the issues encountered by them may be a subset of problems that people may experience. On the other side, Li et al. (2011) focused their investigation on participants who wanted to change or maintain their behavior, were used to use self-tracking devices and were also likely very motivated in accomplishing self-monitoring activities. Fritz et al. (2014) involved participants that had been using these kinds of technologies for long periods, reporting that they were not necessarily representative of experiences with these tools. Moreover, all these studies focused on trackers who had clear goals, such as changing behavior, showing a strong motivation in achieving their objectives through PI instruments. Choe et al. (2014), instead, by investigating Quantified Selfers' practices, focused on a specific user group formed by people who are extremely motivated in tracking personal data and are engaged in actively self-experimenting how different variables may have effects on their behavior.

Although Rooksby et al. (2014) enrolled in their study four participants who never (or barely) used an activity tracker before, they mainly focused on people who had used at least one tracker, while the majority were using several tracking and logging technologies. The majority of their participants had already integrated PI tools into their daily lives at the beginning of the study, showing short-term and long-term goals on the basis of which they used trackers and collected data.

Given these premises, previous research can only offer us limited insight into what people with a curiosity about but no previous experience of PI might need in such tools. A research that investigates needs, desires and perceptions of users with no previous experience in PI is then still missing. We want to fill this gap by conducting a diary study with them.

3. DIARY STUDY: METHOD

To study how users use and perceive PI tools a *diary study* has been conducted. Diary studies are a longitudinal technique to investigate user behaviors in the field reducing the effects of the observer: they diverge from other field study methods as the observer is far from participants and participants control the timing of recording and the aspects of the recorded behavior (Carter and Mankoff, 2005). A diary study commonly uses a paper diary or a notebook, in which users record the time of an event, information about the context and the significance of episodes that have occurred to them, along with thoughts and feelings associated to those episodes. Diary studies are useful as they allow to capture information in contexts that would be difficult to directly observe for a researcher due to social or physical reasons, "potentially leading to more personal accounts and natural behaviour" (Church and Smyth, 2009). This methodology has been used for studying the usage of a variety of technologies in Human-Computer Interaction field, such as SMS (Grinter and Eldridge, 2003), multitasking tools (Czerwinski et al., 2004) and mobile devices (Sohn et al., 2008)

Conducting a diary study, we realize that there would be imperfections in data gathering, as the act of observing and recording tends to interrupt the flow of daily activities. Furthermore, diary studies suffer the disadvantage of possible missing data, since participants are selective in reporting (e.g. they could think that some events are not important while they are). Despite these factors, we thought that a diary study would be the most effective methodology to capture contextual data and highlight user needs and barriers in using PI tools: we needed a technique that could collect data in situations where it would be impossible to have direct observation (e.g. at night in the user bedroom), capturing individuals' personal descriptions and perceptions in an everyday context of use.

3.1 Sample

We involved a diverse group of people from our city, as the physical proximity of the participants was important in conducting contextual interviews that we carried out at the end of the study. Participants were recruited through a variety of approaches including snowball sampling and recruiting emails. Fourteen participants took part in our research, 8 female and 6 male. Other studies with similar design and purposes in HCI field have adopted a similar sample size (e.g. Li (2012), Grinter and Eldridge (2003), Czerwinski et al. (2004)) However, the final decision of settling for 14 participants came when we realized that additional data would not have produced substantial new insights for the purposes of our research, following a data saturation criterion (Bowen, 2008). Background information was collected through a preliminary interview.

The age of participants ranged between 19 and 50, with an average of 31,9 (SD: 10,1). As for professions, our participants comprised undergraduate students, PhD students, post-doctoral researchers, a psychologist, a commercial operator, a web designer, a software developer, an e-commerce sales manager, and a lawyer. All participants owned a smartphone: 7 out of 14 used regularly applications and mobile internet access on their phone; the other seven used their smartphones mainly for calling or messaging. All participants were open to technology. However, all except four were not focused on technology (i.e. they did not work in a technology company or study technological disciplines). Only two were moderately adept in data analysis or statistics (the software developer and the psychologist). None of the participants were obese. While two (one man and one woman) were moderately concerned about weight none of the participants were dieting. There were no serious runners, or professional sportsman, among the participants, but five regularly exercised by running once a week (two), or by walking around thirty minutes a day (three). Four (of which two also exercised) were following, without striving too much, a healthy lifestyle by searching for natural foods, by being health conscious, or by paying attention to the factors that could negatively impact their wellness (e.g. too much pollution, excessive stress, etc.). Four participants felt somehow the need of doing more physical activity, or following a healthier lifestyle, but they did not believe to have sufficient time and motivation for changing something in their life. Anyhow they did not care too much of these matters. The remaining participants were not concerned about health or doing physical activity.

In regards to the act of collecting their personal data through technological means, all participants shared the following attributes:

- they did not have any prior experience neither with PI tools, nor with the act of collecting their personal information (differently from the users of the Li et al. 2010, 2011, Choe et al., 2014, Fritz et al., 2014 and Rooksby et al., 2014 studies);
- they did not have any special need or specific goal for recording their personal data, such as a chronic disease or a desire for changing their behavior (e.g. quitting smoking) (differently from the users of the Li et al. 2010, 2011, Choe et al., 2014, Fritz et al., 2014 studies and from the majority of the Rooksby et al.'s users, who were dieting or serious runners);
- they did not have a deep knowledge of self-tracking instruments and, by and large, of wearable and ubiquitous technologies and they did not seek such a knowledge (differently from the Choe et al.'s users and the majority of participants that were recruited in Li et al. 2010, 2011 studies);
- they had a positive predisposition toward the idea of gaining a better knowledge about some aspects of their behavior and of their daily life;
- they showed a sense of curiosity toward Personal Informatics systems.

Six participants reported that they have already heard about this kind of tools, they were interested in them but they have never had the occasion to try them. Four participants wanted to buy a wearable device in the last year, but then procrastinated the purchase, as they could not concretely figure out how this could be integrated in their everyday life. The remaining participants showed a superficial knowledge of what wearables are, and did not know about tools for self-tracking. However, they were aware that some people monitor their behavior.

When PI tools were assigned to them they all had a positive attitude in trying to understand how these technologies could be useful in their everyday life. Although they did not have specific goals in mind, they were curious and interested in collecting and exploring their personal data. We suggest that these participants, in their unique ways, can provide good cases of people that could adopt PI tools in the near future,

We will refer to these users as *naïve users*, meaning a class of new, potential users of PI technologies, who lack experience and knowledge of self-tracking, but who are curious and interested in it, in contrast to the *experienced users* involved in the research cited above. Participants were not compensated for their participation.

3.2 Procedure

We split participants into two groups of seven users using two different types of Personal Informatics tools: the first group had to use a wearable device, the Jawbone Up bracelet, recording data for physical activity, sleep, food and mood; the second group had to use three PI applications on their smartphones, collecting data for dreams (Dreamboard), sleep (SleepBot), movements (Moves), locations (Foursquare), physical activity, food and calories (MyfitnessPal), daily tasks (Daytum), or mood (Expereal, T2 Mood Tracker). Jawbone Up bracelet automatically tracks physical activity and sleep patterns by detecting the movements of the wrist. Food and mood, instead, need to be manually inserted by the user through a free text field and a graphical interface (a measurement scale). Dreamboard allows users to keep track of their dreams by combining simple icons and free text fields. SleepBot automatically detects sleep patterns by tracking sounds and movements at night, while Moves exploits the sensors embedded in the smartphone to track the number of steps. Foursquare records locations by requiring users to make a check-in, while MyFitnessPal provides users with a food journal, requiring them to manually insert the food intakes or to scan the barcodes of the food containers. Daytum is an app for collecting and displaying daily activities and events that should manually inserted. Expereal and T2 Mood Tracker enable the tracking of mood through a graphical interface (measurement scales).

The data tracked and the tool(s) used by each participant are shown in Table 2. Participants were assigned to the different groups on the basis of their personal interest and curiosity in the parameters to be tracked, and motivations for wearing on not wearing the device (e.g. aesthetics, suitability for their everyday habits, etc.).

For example, some participants showed the willingness to know something about a behavior not traceable by the Jawbone Up (e.g. dreams or locations): so they were assigned to the second group. Others did not like at all the design of the bracelet: therefore, they were assigned to the second group too. Other participants expressed a strong curiosity in trying a wearable device that could automatically detect their sleep and physical activity: they were then allocated to the first group.

In regard to the first group, we chose the Jawbone Up bracelet after carrying out a desk analysis and excluding usability problems. We preferred it to other wearable devices because of its relative widespread diffusion, its unobtrusiveness and its ability to gather different kinds of behavioral data (physical activity, sleep, food, mood) through a unique solution: FuelBand, for example, is more addressed to physical activity and doesn't allow to track users' sleep, while FitBit is more obtrusive for its larger dimensions. As for the second group, we selected the applications following these criteria: i) heterogeneity of the behaviors/parameters to be tracked and domains addressed; ii) balance between the self-tracking modalities (automated vs. self-reported); iii) complexity of the data reported. For example, as for the second criterion we chose Moves and Foursquare because the first one automatically detects movements, while the second one asks the user to report her location by doing a check-in: although the latter was not originally designed for self-tracking purposes, it may be used also for keeping track of the places visited (Rooksby et al., 2014). As regards to the third criterion, we selected both Expereal and T2 Mood Tracker because the first one allows users to report mood by simply inserting a value (from 1 to 10) through a colorful wheel; while the second one allows users to specify their mood along different dimensions, resulting in a more complex information recorded. We kept out PI tools that track data for very specific needs, such as those focused exclusively on health issues (e.g. trackers for symptoms or chronic illness) because they were not suitable for our participants (who did not have such specific needs). Moreover, the applications were selected only after excluding usability problems through a heuristic analysis.

These tools were then assigned on the basis of the following criteria: i) interest of the participant in the target behavior / parameter to be tracked; ii) balance between automatically recorded data and self-reported data among the group. For example, we assigned to each participant at least one app that could automatically track data, in order to lighten the burden of self-reporting.

ID	Data Tracked	Tools or Devices	Period of usage
P1	Food, Mood, Physical activities, Sleep	Jawbone Up	1 Month
P2	Food, Mood, Physical activities, Sleep	Jawbone Up	2 weeks
P3	Food, Mood, Physical activities, Sleep	Jawbone Up	10 days
P4	Food, Mood, Physical activities, Sleep	Jawbone Up	3 weeks
P5	Food, Mood, Physical activities, Sleep	Jawbone Up	2 weeks
P6	Food, Mood, Physical activities, Sleep	Jawbone Up	3 weeks
P7	Food, Mood, Physical activities, Sleep	Jawbone Up	10 days
P8	Run, Dreams, Tasks	Nike+, Dreamboard, Daytum	2 weeks
P9	Movements, Locations, Mood	Moves, Foursquare, T2 Mood Tracker	2 weeks
P10	Movements, Run, Mood	Moves, Nike+, Expereal	3 weeks
P11	Physical activities, Movements	MyFitnessPal, Moves, Foursquare	3 weeks
P12	Mood, Sleep, Physical activities	Expereal, Sleepbot, MyFitnessPal	3 weeks
P13	Tasks, Locations, Sleep	Daytum, Foursquare, SleepBot	10 days
P14	Mood, Movements, Dreams	T2 Mood Tracker, Moves, Dreamboard	2 weeks

Table 2. Sample

Participants of the first group ("wearable device") were provided with a Jawbone Up bracelet and had to download the Up application on their smartphones. They were asked to wear the device during the study, night and day. They could self-report in the application both their mood and their food consumption as often as they wished, while physical activities and sleep behaviors were automatically tracked by the device. They were also given the possibility to set goals related to sleep and physical activity and monitor their progression.

Users of the second group ("mobile applications") had to download the assigned apps. As for the applications that relied on self-reporting (e.g., DreamBoard, MyFitnessPal), participants were suggested to fill in regularly the data related to the occurrence of a significant event (e.g. a dream, a lunch). The applications (e.g., Moves, SleepBot) that automatically recorded participants' behaviors (e.g., movements, sleep) were required to be kept active for the whole duration of the study. Participants of both groups were also invited to screen regularly their daily data.

Assigned tools had to be used for a minimum of ten days by each person, keeping a diary of all their thoughts, perceptions, problems and needs related to their usage. However, depending on their availability and interest, they could choose to carry on the trial and the diary activities up to one month. We gave participants the possibility to choose the duration of the trial, as we aimed at exploring their natural engagement and interest in the instruments provided, like so their sustainability in time in their everyday practices. The different user experiences, which came from the different periods of engagement, were considered a richness rather than a confounding variable, as they allowed us to expand the heterogeneity of the cases investigated. One user decided to extend the study for the whole month, while the average period of engagement was seventeen days.

The researchers gave each participant some basic written instructions and an electronic diary in the form of a Microsoft Word™ file, with tables for each day of the week, and rows for each aspect that we wanted to be recorded. Rows were created for: thoughts related to the daily usage of the tool(s), barriers encountered, meanings associated to the data tracked, reflections triggered by the visualized data, unexpected events happened during the usage of the applications, and effects of the self-tracking activity on behavior.

Participants were suggested to fill in their diary during the day, immediately after an important event has occurred, and in the evening, when they could retrace their daily experience and write their personal thoughts and perceptions. Participants could also take pictures of situations and events that triggered problems or particular insights as for the use of the tools assigned to them. Lacks in completing the diary were equally distributed along the whole period of study (there were no peaks in the last days of the research). Thus, the fading engagement that most users showed during the study (e.g. failing to manually report data in the tools provided) was likely related to a decreasing interest toward the tools, and not toward the study itself, as all of them continued to regularly take notes in the diary even in those days when they forgot to use the instruments assigned. On the day after the completion of the diary, participants were interviewed for approximately one hour. They were required to bring the tools they used and the diary they filled to the interview. The interviews were qualitative and revolved around the usage of the assigned tools. Interviews were audio recorded.

4. DIARY STUDY: RESULTS

We transcribed the recordings of the interviews and we analyzed them together with the diaries. We identified themes using a thematic analysis (Braun and Clarke, 2006). During the analysis we focused on the problems people encountered in using the assigned tools. We classified the results in four categories: *tracking* data, i.e. difficulties encountered in collecting personal data both automatically and through self-reporting; *managing data*, i.e. problems faced in integrating different kinds of data and in exercising control over them; *visualizing data*, i.e. issues found in interacting with numeric and abstract representations of data; *using data*, findings related to the PI tools' perceived utility and their sustainability in the long-term.

Regarding the *tracking of data*, participants perceived the act of self-tracking as a burdensome activity, which did not reward them with strong benefits, such as an effectively and immediately increased self-awareness or an improvement of their daily activities. As a result, participants of the second group often failed to self-report their data in the tools assigned, due to a variety of reasons, such as lack of time or motivation and forgetfulness. On the other side, participants that had to wear the Jawbone Up device, highlighted the problems emerged during its usage (e.g. an aesthetic that is not suitable to all the social contexts, interferences with the manual labors of the everyday activities), more than the improvements that it provided to their everyday life.

As to the *management of data*, participants stressed how the tools provided were not able to integrate the different kinds of behavioral data they were collecting. While participants of the first group highlighted that, despite their presentation in a unique interface, the different data collected by the Jawbone Up were simply juxtaposed, users assigned to the second group described that the data were scattered in different silos, preventing them to find useful correlations and insights. On the other side, participants of both of the groups emphasized the need of a greater control over their data, revealing the importance of privacy issues.

The *visualization modalities* offered by the tools presented different problems, such as an excess of abstraction and the lack of recommendations and synthesis. Participants showed to be refractory to numbers and graphs, preferring more concrete and intuitive ways of displaying quantitative data. Moreover, they desired reports and suggestions capable of summarizing important data and highlighting useful insights for their daily life.

Finally, as regards to the *usage of data*, participants found the tools scarcely suitable for their everyday situations. They could not find concrete goals on the basis of which using the information collected, mainly because their engagement started soon to fade away. By perceiving a low cost/benefit relationship, participants were not motivated to integrate these instruments in their everyday practices and objectives, rapidly losing the interest and curiosity they had at the beginning of the study. Although they were interested in knowing something about themselves, they were not disposed to continue to track using the tools provided.

We will present these results in the following sections. However, before extensively describing these four main themes, we outline, in Table 3, a picture of how all these categories of problems evolved over time. The table contains quotes that come exclusively from the participants' diaries, highlighting how they experienced barriers during the four weeks of the study. All the quotes are meant to exemplify how a certain issue was perceived in a particular point in time through the participants' own words. It emerges how the tools assigned did not offer sufficient advantages for engaging them even in the short term: negative thoughts, feelings and judgments were more frequent after the first week of the trial, showing that these tools were unable to maintain the level of involvement at a certain stage. Most of the participants decided to quit after 2 or 3 weeks from the beginning of the study.

First Group	**First Week**	**Second Week**	**Third /Fourth Week**
Tracking	"I reported the mood three times today" P4; "Blamed for using the phone during the lunch" P2	"I don't remember how I was in the afternoon. I inserted good but I'm not sure" P1; "I stopped to insert the mood. I can't make myself remember it" P4	"I forgot to wear the bracelet today and yesterday. It doesn't seem useful to me anymore" P4
Managing	"No control in sharing data" P2; "All these data are not merged. They are put side by side" P1	"When I wake up late, I have always a bad mood. Why?" P5; "Making less than 6000 steps per day is making me sleep less than 5 hours/night? Or is it the contrary?" P7	"Who is the owner of my data?" P5; "I would like to know how Jawbone is using my data" P3
Visualizing	"Beautiful graphs, but I prefer evocative representations (avatars?)" P2; "It would be smart to have some graphical reports, in the form of	"Jawbone should give me advices, like: you are going to sleep bad because you didn't do enough physical activity this day, don't drink tea or	"I'm continuing to track myself, but I'm not visualizing the data anymore. It's always the same thing: bars, bars, bars" P5

	images" P4	coffee this evening" P1	
Using	"What have I done the last Monday at four?" P3	"It can be used only for a short time to understand my problem" P3; "Not useful for remembering what I've done" P2	"Too many practical problems. Third week and it seems to me as it were the third month: this Friday I'll quit" P3
Second Group	**First Week**	**Second Week**	**Third Week/Fourth Week**
Tracking	"Writing just after waking up is not for me. It is too complex when you are sleepy" P8; "Moves is perfect you don't have to do anything. It only consumes a little bit battery" P10	"Too cumbersome to insert data every day. I dropped" P12; "T2 canceled. Very annoying alarms" P14	"This week I didn't use any app. I didn't have time to insert all those data" P12; "I inserted my food for two weeks, but now I'm tired. I won't do it again" P11
Managing	"I need one tool, not three different apps" P10; "I'd like to send my data only to one person not to public them on Fb" P8	"Impossible to keep traces of all these different data without using pen and paper. It's not for me" P13; "Moves is wonderful. I want all my data in this app" P14	"I don't feel safe by having my data somewhere in the net" P10
Visualizing	"Too many numbers. It's quite confusing" P9; "It would be beautiful to read these data as they were stories" P8	"Didn't visualize the data today. They confuse me instead of giving me new ideas" P8	"I like to see how much I moved in the last two weeks. But so what? It lacks something" P11
Using	"Useful if you have a chronic disease" P10; "Today I understood how much are 8000 steps. That's great" P11	"It's nerds' stuff. It's not for me" P8; "Useful until I understand why I'm sleeping so bad" P13	"Interesting but I won't use them after finishing this study" P10; "I will use only moves. It makes me

			remember the things happened" P12

Table 3. Diary Results (sample quotes from the participants' diaries in the cells)

We will now move to the description of the four themes found during the data analysis, focusing on the problems that participants encountered, in order to highlight the critical points to be solved for making PI tools more suitable for people's everyday needs. Most of the issues found were experienced by almost all the participants, highlighting, on the one side, that their needs and desires were somehow homogeneous, and, on the other side, that the same problems recurred in different tools, apart from the specific data they tracked and the particular interfaces they had. The findings presented refer to both the data collected through the diaries and those gathered during the interviews, as they did not show any significant difference or contrast that could justify a separate exposition. Nevertheless, in the subsections above, all the quotes come from the interviews: as during the interviews participants had the possibility to also comment the data they gathered through the diaries, the quotes we collected in such occasion better exemplify both of the sources.

4.1 Tracking Data

Issues in manually collecting data

People often failed to self-report their data, which required their active role in the gathering process. Participants reported different problems with this task. First, people did not remember to insert data. For example, P10, after three days from the beginning of the study, started to forget to track the majority of his emotional states. P1 remembered to insert mood mainly in the evening, when she explored her daily data. However, often it was too late to recollect her exact emotional states experienced during the day: this entailed a more or less casual reporting of what she felt like. P13 remembered to insert data only after events had happened. However, at that time she lacked motivation and willingness and rarely went through self-reporting: "*it was gone already, and I said to me I'll do the next time*". P8 lacked time and willingness in reporting her dreams just after she woke up. Inserting complex textual information in a mobile phone in the early morning did not get along with her daily habits. By waiting a few hours, it was too late, and the dreams were gone.

Participants experienced problems also when the data to be reported required a minimum effort, for example using the measurement scales adopted by Jawbone Up and Expereal. Although P1 and P3 appreciated the intuitiveness of the task related to the insertion of mood, P2, P4, P6, P10 and P12 reported that signaling mood only along one axis (positive-negative) oversimplified their emotions: they did not recognize their psychological states in the dimension provided. P4, for example, reported that the degrees of smile used by Up application were "*somewhat fuzzy. They can be useful for expressing emotions while chatting but not for tracking psychological states in order to rely on accurate data*". P2 recognized that "*emotions are often contradictory and applications should provide the possibility to select different emotions at the same time, qualified by names and intensity*". On the other hand, providing too many dimensions, as done by T2 Mood Tracker, was perceived as confusing and annoying. P14, for example, did not understand the differences between the various axes provided by the tool (e.g. worried, pressured, tense, irritable, anxious), noting that "*reporting emotions should be intuitive such as experiencing them*". These elements highlight how self-reporting activities can also arise issues related to the complexity of the data to be reported: adding too many details can discourage users in manually insert the information, while simplifying too much the process of recording can impoverish the data, making it useless when retrieved in a later time.

Manually collecting information was also perceived as an out of place task that did not fit well with everyday practices of participants. P12 considered that reporting food intakes was an activity that could not be easily performed during the meals, since it was stigmatized by her dining companions (e.g. family, friends, colleagues). P2 reported that she felt strange to interrupt every meal to put data in a mobile application. P3 and P7 agreed that in many everyday situations it was hard to gain social acceptance when stopping social activities to deal with an electronic device, especially when the actions required for self-reporting took quite a long time. Other participants, instead, found self-reporting to be a distraction from their actual main task. P6, for example, stated that inserting his mood, through the UP mobile application, required to take out the phone, open the app, select the mood and close the application: this procedure diverted him from his current thoughts and actions and for this reason was soon abandoned.

Thus, while gathering information immediately after an event had occurred was perceived as an interruption or an interfering activity, collecting data after a while showed not to be a viable option either because of forgetfulness or lack of motivation. As a result, 13 out of 14 participants agreed that manually collecting data was too cumbersome to carry on in daily practices. Regularly self-reporting was interpreted as undertaking a new habit with too high costs. Meanings most frequently associated to this in the diaries were *cumbersome*, *annoying*, *unsuitable*, *complex*, *stressful*.

Issues in automatically collecting data

While participants experienced many problems in manually collecting data, automatic self-tracking features, provided by Jawbone Up, Moves and SleepBot, were welcomed by almost all the participants. Participants were impressed of the potentialities of the tools provided, reporting a sense of wonder for their capabilities. P10, for example, was struck by the fact that Moves was able to recognize when he was walking, running and cycling, automatically differentiating the three activities; while for P2 it was hard to believe that from the movements of her wrist it was possible to detect the different stages of her sleep.

However, participants of the first group reported that carrying a wearable device like Jawbone Up may give way to many practical problems. Although comfortable, P1 stated that "*it is unimaginable to wear a bracelet like that for a person that does a manual labor*". P3 reported that it interfered in the daily manage with her child, as it could be harmful when she held her. P6 observed that it damaged the liner of his jackets and scratched his face during sleep movements. P5 had an annoying sensation of pressure on the wrist. P7 believed that it blocked the flow of blood during the night. On the other side, participants sometimes forgot to charge their bracelet, making it useless the day after: P6 reported that "*I realized that it was low on battery only when I was out and for the rest of the day it didn't track anything*". Others, instead, reported their forgetfulness in wearing the device, especially after a shower or after a manual labor. P3 for example stated that "*Usually after a shower I forgot to wear it for three or four hours, one time it's happened that I forgot to wear it for the whole day*".

Moreover, participants highlighted how this kind of devices bump into social and "fashion" issues, like their integration with one's personal style: while P4 noted that the bracelet was not suitable for all daily contexts, due its appearance, P1 and P2 agreed that devices that need to be worn everyday and everywhere do not fit with the need of changing garments according to different days and situations. Especially female participants emphasized the importance of aesthetical matters when dealing with this kind of objects. If P5 thought that its minimal aesthetics perfectly fit with her "*casual dressing style*" of her everyday life, P3, instead, pointed that the light blue of her Jawbone could not be appropriate for a nightdress or for formal occasions, while it could be suitable for a sport dress when going out for a run. All these participants highlighted the social valence of objects that are constantly under the eyes of other people and should reflect particular and also temporary climates of their owners.

Members of the second group encountered different problems. P9, P11, P14 complained that their phones' batteries were drained by Moves so much so that P9 and P11 reported that often they needed to close the application in order to preserve the functionality of their phone. P14, instead, stressed that she was inclined to have less battery for using a useful app as Moves, but the charge issue should be solved in its next release for exploiting all its potentialities. Also P12 and P13 stressed that they had to keep their phones connected to their chargers during the night to avoid the battery drain caused by SleepBot: the first day they found their device out of power just before they were going out for work. They stressed another point that they considered important: their phones fell from the bed during the first nights, forcing them to put them under the pillow away from the edge of the bed. Although this was a more secure place for the phone, they did not feel safe in having it so close to their heads all night long.

Another critical point was related to the perceived accuracy, and thus reliability, of the gathered data. Participants doubted of their trustworthiness when they noted a discrepancy between some data automatically collected by the tools and what they believed about their behaviors. This matter was important especially for participants that regularly exercised. When they had to choose whether to rely on their memories or on the data collected, they always preferred the first option, calling into question the capability of the overall system to correctly measure their behavior. P6 and P7, for example, found that the level of activity tracked by Jawbone Up did not reflect their beliefs about their daily movements. This issue generated doubts about the sleep data too, causing a rapid disillusion about the trustworthiness of the system. However, this reaction was mitigated when referred to the self-reported data, in which participants had the responsibility of their accuracy. P11 explained this attitude: "*When I'm the responsible for reporting my data, I may know where and when I was not accurate, so discrepancies between what the system displays and what I know about myself is not a big deal. But when the system fails in collecting my data, I may not know where and when it got wrong, so how can I rely on it?*". Reliability, then, was a feature that participants attributed more to the responsible of the data gathering (the system or themselves) than to the gathered data: and although they were indulgent about themselves, they were inclined to withdraw their whole confidence in the system at its first signs of hesitation.

Scarce help provided by tools

Although some issues in collecting data occurred because of the users, tools were of no true help in speeding up the logging activity and reducing the cost of self-monitoring. Reminders mostly bothered participants, especially the professionals and those who were more technology conscious, rather than helping them not to forget to insert data. By using reminders, these apps were interpreted as a potential source of additional "noise" for people that were already drowning in notifications for their daily work. The messages that Expereal sent to P10 regularly, for example, were perceived as an annoying interruption from his activities and were ignored. P14 canceled T2 Mood Tracker application before the end of the study because she could not stand its daily alerts. P9 reported that every time he felt his phone vibrating, he thought that it was a message: but after extracting it from his pocket he discovered that it was a T2's reminder, making him quite upset.

Other features, aimed at facilitating the insertion of data, raised many complaints too. Pre-sets for food, for example, were perceived as culturally determined, specifically addressed to American people. P3 wanted to report her food intakes but did not find what she usually ate in the list of the Up application. P12 tried to create new food items and recipes with MyFitnessPal but this task required too much time and was scarcely engaging: after few attempts she decided to abandon it. The scan barcode feature, on the other side, seemed very useful at first sight by P11, but he soon realized that it could work only for a limited set of items and that complex foods would not result as the sum of scans of their parts.

4.2 Managing Data

Lack of integration

People found that they were unable to see the whole picture of the information they gathered. Tools provided few possibilities to integrate different kinds of behavioral data. P8 and P9 reported that apps were not helpful in connecting the various events, while P11 and P12 wanted an application able to mix the data coming from the different tools they used. By being scattered in different locations, data collected prevented participants to find useful correlations among them. Actually, people wanted to find relations between different kinds of information, such as co-variation and causation. However, as P12 stressed, they were not supported by the tools: she wanted, for example, to see whether her meals had some influences on her mood and her sleep, but the dispersion of these data impeded her to infer useful insights. On the same way, P2 desired to know how her sleep was connected with her daily physical activities: although she had all this information stored in the Up application, she was able to see only co-variations between them, without understanding the reasons behind them.

The main point is that participants did not proceed in self-experimentation to verify the foundation of the co-variation found for translating them in valid causations. They showed not to have sufficient motivation and time to manage their data, reflect on them and act to discover useful connections. Instead, they wanted the tools to provide them with the most likely relationships, suggesting valid causalities. P1, for example, stated that she was not able to decide which of the co-variations between food intakes, sleep, physical activities and emotional states could be relevant for her, wishing for a more active role of the tool in providing important connections. P6 desired more intelligence in tools: "*I do not have the knowledge for saying this caused that, I can see a co-variation but actually it could be only a mirage, like saying I'm sad every time it is raining in Brazil. The tool should provide me valid causations that rely on scientific bases*". People wanted to maximize the cost-benefit relationship, transferring to tools the burden of finding causations and expecting sufficient intelligence in the systems to validate them.

Lack of control

Although younger participants were slightly less concerned about privacy matters, information tracked through PI tools were perceived as extremely private, and almost all participants (except P9) agreed that they would not share them on social networks or expose them in public. However, some of them showed the desire to compare some of their data with those of significant others. P3, for example, wanted to match her sleep patterns with those of her husband, while P5 liked to share only a subset of them with her boyfriend. P4 said that "*I would have liked to send my physical activity data to my sister, but the only thing that I could do was to share them on Facebook or Twitter*". As total publicity was seen as a threat for participants' privacy, sharing specific data through private communication channels or among a strict group of well-known people was perceived as the unique viable option for disclosing their objectives and results to others.

The keyword stressed by almost all the participants was referred to the control they should exert on their personal information. P1, for example, stated that these tools should allow users to decide which data should be stored, deleted or made public in every moment. P9, instead, call the security of the storage mechanisms into question, stressing that all these private data, distributed in different places, could be soon missed in the web, making impossible for the individual to retrace them. Losing control was thus the most dominant fear among the participants: they expressed a feeling of reduced trustworthiness toward the service providers, as the possibilities for conserving, sharing and deleting their information were not considered under their decision. Moreover, the tools used did not provide clear information about who, where and how long all these data would be stored, leaving them a feeling of uncertainty that further questioned their trust in this kind of technologies.

4.3 Visualizing Data

Excess of abstraction

Although several participants (4 out of 14) welcomed the possibility of displaying information through graphs and stats, the majority of them reported that data visualizations provided by the tools were too abstract and removed from what they expected. P3 did not recognize herself in the data displayed. She was not able to establish an emotional connection with them, which would have helped her give them more importance "*These information are somehow cold, how can I reflect myself in there?*". P1, P4, P8 and P12 agreed that abstract representations are addressed to people with a passion in exploring numeric information. They, instead, preferred more concrete visualizations in which mirror and recognize themselves. Developing an emotional link between users and their data was a way to become involved in exploring and giving meaning to what the instrument gathered and presented back.

On the other side, the immediacy of the representation was considered an essential requirement for facilitating the comprehension of the information collected. P1, for example, wanted an impressionistic image able to provide, in a glance, key information and their relationship in a holistic way. For P2 the tool should have fed back a global representation of the user's lifestyle, in a metaphor that could portray intuitively and synthetically what she have gathered.

As a result, over half of our participants expressed a lack of engagement in visualizing and exploring their data. P5, P8 and P13, for example, in the last days of diary study did not display their data any more. The inability of the tools to provide an immediate engagement reduced the participants initial curiosity and interest, preventing them to continue the reflection on data and consequently on themselves. Others were instead affected by a shallow attitude, determined by the scarce clarity of the information displayed. P2, P3 and P6, for example, reported to be surprised that their deep sleep was not continuous and that their light sleep lasted so much time. However, when researchers asked them why deep sleep should be so important for them, they were not able to answer: they associated a positive valence to deep sleep and a negative valence to the light one, without trying to further understand the meanings of these two terms. We encountered this superficial tendency in other participants too. People showed a lack of motivation in navigating their data, not going in deep in understanding them: they quickly visualized some information, attempted some correlations, but were stopped at the first difficulties.

Lack of synthesis and suggestions

Most participants desired brief reports highlighting the most important data of the day. They often found the visualized data hard to understand, requiring a high cognitive load and long time to find thought provoking insights on their behaviors and their activities. P8, for example, stated that data should be visualized both in graphical form and in narrative form, as statements connecting different kinds of data with the events happened during the day. Most of the participants suggested that reports should be connected with recommendations on how the data could be useful for their daily life. P6 thought that suggestions were fundamental to address a change in his behavior. They represent a form of external judgment that can motivate people to improve themselves: "*I know that I have to do more physical activity, but seeing it on the display of my phone is like an external evaluation, it pushes me more to change my habits*". As stated by P2, P9 and P10, these reports should be in a linguistic form, helping users understand their faults and find the best solution for them, "*like a good physician or therapist should do with his patients*". However, they also noted that transforming numeric data in natural language could not be sufficient: they should also involve users in the process of reading, as a series of disconnected phrases could be annoying as the actual representations.

Tools did not show sufficient “intelligence” for advising participants in how to exploit the gathered data for their wellbeing. P5, during the diary study, wanted to use the data tracked to change her daily physical activity and improve the quality of her sleep. Nevertheless, she reported that she was not able to address her efforts correctly, as the application did not provide any suggestion: “*Up should know me, since it has all my data, and should give me suggestions tailored on my physical and demographics parameters, such my age, gender, weight and on my daily habits. I don’t have scientific or medical knowledge to help myself and I have no time for exploring on my own how I should improve my behaviors in the right way*”. As long as participants wanted useful insights applicable to their everyday lives, personalization was reported as a main wish: since the tools had a lot of data regarding user specific behaviors, it was natural for people to imagine that they could provide services targeted to their specific needs. For P11 suggestions should show different alternatives, tailored on the basis of the user’s characteristics, while P3 wanted the system to recommend goals too, on the basis of the knowledge that the system should have about her.

4.4 Using Data

Perception of utility

People experienced a sense of initial curiosity for PI tools and during the study they tried to imagine how these technologies could be useful for them. Several participants thought that they would use these tools to track what happened to them and recollect their memories even after a long period of time. This kind of use was considered a way “*to preserve all those things that we don’t notice during our life, but that could become important later*”, as stated by P14, or for “*re-experiencing past episodes that we believed lost*”, as wished by P8. Continuously monitoring one’s own behavior, then, was interpreted as a means to enhance individuals’ memory. However, reliving the experiences connected with the data collected was considered also essential for giving them meaning: without memories, personal data resemble to void shells, useless for gaining insights on users’ everyday lives. P2, for example, reported how it was hard to recollect what she was doing three days before, looking at Jawbone Up’s graphs: “*Here I see a peak in my physical activity at 12 o’clock. But sincerely I can’t remember now what I was doing in that moment, this means nothing to me*”.

Participants, like P4 and P7, pointed out that the timeline, used for example by the UP app, was an important axis, but not sufficient, alone, to connect the information collected to a meaningful background. Instead, participants that used Moves were helped in recollecting the memories of their daily activities by the traces of their spatial movements: space was then used as a frame through which remember what happened during the day. Looking at Moves, P11, for example, was able to recall on the one side the distances covered, on the other one the people met and the accomplished tasks. Moreover, P14 desired to integrate Moves with the data related to her mood, since this would allow her to connect her emotional states to the places visited, and, through them, remember what happened to her during the day. These suggestions stressed how spatial cues were essential in triggering the process of remembering.

The majority of participants thought that PI applications and devices should be also useful for modifying their behavior: although not enough, this kind of technology could enhance their motivation and willingness to undergo change. P11, for example, said: "*I didn't know how much eight thousands steps could be. Now that I get the measure I can set myself to that goal*". However, this attitude contained manifold and ambiguous aspects too. Several participants, especially those who were not already exercising or following a healthy lifestyle, said that this kind of technologies would only be useful for people affected by a chronic disease, or strongly motivated to change an unhealthy behavior. This point highlights how these tools were not perceived as helpful by people that have an ambivalent attitude toward change, who may be aware of the need to change, but, at the same time, believe not to have the sufficient resources for carrying it out. Another perspective was expressed by P8 and P13 that thought that PI tools were suitable only for "*nerds*" (i.e., people enthusiastic for new technologies). It is interesting to highlight that participants thought about all these kinds of possible users as "others than themselves". They did not identify with individuals strongly interested and motivated in searching for and better knowing the factors that may have influence on their behavior. Instead, they were interested in self-monitoring activities that did not require hard efforts in order to be accomplished with success.

Long-term usage

Over half of our participants stated that this kind of tools could be useful for changing one's own behavior toward healthier lifestyles. However, P3 believed that tools like Up could be useful only over a limited period of time. She would use it only for the time needed to trace a baseline of her current behaviors: "*Once understood where I'm wrong, I don't see any reason for wearing it again. Maybe if one day I will change my* *habits, it could be useful to me again*". P5 agreed on this pattern of usage: PI tools are useful only until users understand the specific behaviors that they want to investigate, or until they address the problem that may affect their daily life.

Besides, their suitability for a long-term use was rated very low. People perceived a low cost/benefit relationship: the burden of tracking data was too high in comparison of what the tools were able to provide them with. The majority of participants experienced a sense of curiosity at the beginning of the study, which rapidly decreased as the days passed, together with the willingness and perseverance in tracking their own behaviors. "*It was interesting for the first times, but what else is there?… after few days it became boring reporting my information every day and using that application for seeing more or less always the same things, the same graphs, that's not much*" said P5. Although several participants thought that using these tools for a long time could impact their self-awareness and consequently drive some kind of change, they also expressed that complying in such a burdensome activity, without immediate gratifications, was not suitable for their lifestyles.

5. DISCUSSION

Our results build on top and extend previous research, confirming and contrasting with studies addressed to investigate how experienced users use PI tools. At the beginning of our diary study, our users differed from those recruited in previous works on commercial PI tools usage (Li et al., 2010, 2011; Choe et al., 2014; Fritz et al., 2014; Rooksby et al., 2014) on the basis of the following main accounts: i) they did not already integrate PI tools in their everyday life; ii) they did not have a clear understandings of how these tools could be exploited for their own situated purposes. However they shared an interest and curiosity in personal data and in knowing better something related on their own everyday life. Additionally, our participants also differed from those recruited in many of the trials carried out for assessing novel PI prototypes (e.g. Consolvo et al., 2008; Tosco et al., 2006; Rooksby et al., 2015). Assessing a prototype frames the trial differently, because users know that they might contribute to better the system under evaluation. Moreover, our participants did not have any specific goal in mind at the start of our study (while the users recruited e.g. for the Ubifit Garden's trial (Consolvo et al., 2008) and the Chick Clique's trial (Tosco et al., 2006) shared behavior change goals, and the participants of the Pass the Ball's evaluation (Rooksby et al., 2015) aimed at winning the game under the frame of which they were participating).

During the diary study and the interviews other differences and similarities emerged. Some of the barriers found in our research actually strengthen the previous findings, showing that some issues are widely recognized as problems (Table 4). The burden of the self-tracking activities, for example, has been highlighted by the Li et al.'s (2010) users as well: when a tool did not satisfy their information needs, they often switched to another tool. The Quantified-Selfers' themselves, investigated by Choe et al. (2014), showed a great tracking fatigue, which pushed them to search new ways to alleviate it, such as making the data collection automatic or lowering data granularity by developing personalized tools. Although encountering more or less the same barriers of experienced users, naïve users expressed a more submissive attitude toward them. Obviously they could not try to build their own instrument, but actually they were not even in search for a solution, but for a way for escaping from a laborious task.

Also as to the *integration* among different sources and kinds of data, naïve and experienced users shared the same difficulties: having the data scattered among a variety of silos makes the interpretative process hard. However, while experienced users, such as those of Li et al. (2011), managed to explore their data by using paper graphs or by reviewing their logs, naïve users requested a ready-to-use solution, showing a rather weak motivation in working on their data. In fact, one of the characteristics of experienced users is their motivation in self-experimenting, i.e. their desire to discover correlation among various information that may affect their behaviors. Quantified Selfers, for example, usually generate hypotheses to test, by making careful observations of previous behavioral patterns or individual needs, and sometimes by using control conditions, triangulation of data and the experience sampling method to design more rigorous experiments (Choe et al., 2014). Instead, naïve users expect a greater intelligence in the tool and transfer the work needed for finding correlations and even causations to the instrument, as they do not think to have the required competences for formulating hypotheses and validating them through self-experimentation.

Moreover, in relation of control, concerns about *privacy* are shared among the two different kinds of users, although to a different degree. For example, the main concerns of experienced users are related to having personal data uploaded to a server or making these information visible to others (Fritz et al., 2014). Naïve users express a deeper fear and a stronger request to have a complete control on their data, likely driven by a minor knowledge of the technological and "economical" issues that lie behind the ownership of the data they generated. Nevertheless the unwillingness to share personal information to generic social networks, such as Facebook and Twitter, is reported by both the experienced (Rooksby et al., 2014) and the naïve users, although the experienced ones seem more inclined to compare their data with friends within the community of the tool they used with the purpose of competition (Fritz et al., 2014)

If these aspects show common *desiderata* related to the tracking and management of personal data by experienced and inexperienced users, we can find that naïve and experienced users have different needs and wishes related to the *visualization modalities*. Fritz et al. (2014) highlighted how experienced users tend to focus on numerical goals and numerical data, suggesting that they pay more attention to numbers rather than on the activities they represent. Instead, naïve users are not very familiar with the visualization of quantitative data provided by PI tools. They prefer more impressionistic representations and suggestions targeted on their habits and behaviors. Nevertheless, data visualization and interpretation is a key hurdle even for the most experienced users. Choe et al. (2014) noted that Quantified Selfers simplify their tracking strategy when there is no easy way to analyze and interpret data. The learning curve in creating the most appropriate visualization for a given data type is very steep also for them, and only after many attempts they are able to find visualizations that are really helpful in raising their self-awareness. Naïve users, on the other side, are not willing to spend long time in learning how to make the display of their information effective. Immediacy and intuitiveness are their most common requirements, as they prefer the possibility of quickly gaining insights to be immediately used in their daily activities, rather than the opportunities of widely exploring and navigating their data from different points of view.

Naïve users are also less goal-driven than experienced users (Li et al., 2011), as they rarely think to have specific objectives to reach through the use of self-tracking technologies. Rooksby et al. (2014) reported how, for most of their users, tracking was directly related to self-esteem, tied to pride at completing a marathon or body-images problems, aging and mental health. Naïve users, instead, by not having clear how these technologies could help them achieve their personal goals, wish for immediate perceived benefits: they want the tool to propose them personalized goals useful for their daily habits and everyday practices, based on the data it continuously gathers.

Finally, in this study we found how naïve users experience and understand PI tools differently from experienced ones. Experienced users integrate their self-tracking devices deeply into their habits and daily practices, wearing them throughout all the day and taking them off just before sleeping, and describe a strong attachment to them (Fritz et al., 2014). Conversely, tools provided to naïve users were perceived mostly as addressed to technology fanatics or people with special desires, but not so close to their needs and their view on what technology should address. They often forgot to wear the bracelet provided, looking at it more as an inconvenience than a precious object. Its aesthetics, as well as its integration among different social contexts and practices, were considered essential requirements for these kinds of tools to be suitable for their daily lives.

By and large, PI tools were used in a manner previously unclassified by Rooksby et al. (2014): we can refer to this style as "playful tracking". The playfulness of this form of tracking is related to the fact that naïve users did not take the self-tracking activity as seriously as experienced self-trackers do. They "played" with data and experimented the tools' features, primarily trying to figure out what kind of benefits they could gain in interacting with them. This style is characterized by curiosity, exploration, serendipity, and willingness to discover something of unexpected: "*I was looking at how many steps I was taking every day, and I discovered that the most interesting thing was how little I was sleeping. This was really cool… without this bracelet I couldn't ever have figured it out*", said P4, while P8 "*It's fun to see something that you didn't expect. But everything should be much simpler, immediate, otherwise a somehow entertaining activity inevitably becomes a difficult and thus annoying activity*". They see tracking as an opportunity to discover something useful in relation to their everyday lives, but, as long as they do not have clear objectives in their mind, they do not want to invest too much energy and time in this task. For people characterized by this style of tracking, fun and surprise are essential elements to get involved in the use of such technologies. They want to be immediately engaged in an experience that can be enjoyable *per se*, quickly showing how these tools could provide them with useful insights. They use these technologies mainly to satisfy their curiosity to find unexpected aspects of their behaviors, while the raise in self-awareness that could come from this activity is a welcomed side-effect that, however, is not intentionally pursued at the first stages of their experience.

Main issues	**Specific issues**	**Similar findings in previous research with experienced**	**Different findings in previous research with experienced users**

		users	
Tracking Data	Issues in manually collecting data	Li et al. (2010) and Choe et al. (2014) highlighted how experienced users find burdensome tracking their own activities.	Li et al. (2010) showed that experienced users switch to another tool when they are not satisfied, while Choe et al. (2014) reported that they can even build their own tool. We showed that naïve users do not struggle to manage the burden of tracking data with "home made" solutions, but they want the tool to simplify this laborious task for them.
	Issues in automatically collecting data	//	Fritz et al. (2014) showed that experienced users integrate PI devices in their life, wearing them almost all the day and night. We showed that naïve users found wearable devices difficult to integrate in their everyday practices.
	Insufficient help provided by tools	//	Naïve users question the helpfulness of the tools, not being satisfied of the actual support, while experienced users, as highlighted by Choe et al. (2014), often try to find a solution by themselves, knowing the limits of the current technologies. Although alerts can be annoying also for experienced users, Li et al. (2011) highlighted how it is possible to force users to interact with their data daily, either by sending reports or reminders. However, naïve users show to scarcely bear this kind of support, stressing that

			other types of help are needed.
Managing Data	Lack of integration	Li et al. (2011) and Choe et al. (2014) reported that experienced users experience difficulties in interpreting their data, by having them scattered among a variety of different tools.	Li et al. (2011) showed that experienced users manage to explore their data together, by using paper graphs or by reviewing their logs. We highlighted that naïve users request a ready-to-use solution, showing a rather weak motivation in working on their data. Choe et al. (2014) reported that experienced users formulate hypotheses to test, while we stressed how naïve users ask for more intelligence in the tools for recognizing useful correlations.
	Lack of control	Fritz et al., 2014 noted how experienced users have many concerns about privacy issues, while Rooksby et al. (2014) noted how they are not inclined to share personal information to generic social networks	We showed that naïve users have a deeper fear of losing their data and are less inclined to compare them with friends for competition purposes, as done by experienced users (Fritz et al., 2014).
Visualizing Data	Excess of abstraction	Choe et al. (2014) highlighted how the learning curve for creating the appropriate visualizations is very steep for experienced users, stressing how they often find difficulties in displaying their data.	Fritz et al. (2014) showed that experienced users focus more on quantitative data, while we noted how naïve users prefer impressionistic and intuitive representations. Choe et al. (2014) outlined how experienced users are willing to try many attempts to find useful visualizations, while naïve users are not.

	Lack of synthesis and suggestions	//	Li et al. (2011) showed that experienced users are goal-driven, while we found that naïve user are less inclined to set their own goals. Differently from experienced users, the usage of these tools is not directly linked with naïve users' self-esteem (Rooksby et al., 2014).
Using Data	Perception of utility	Li et al. (2011) highlighted how these tools are used also for remembering and assisting in behavior change.	Experienced users use PI technologies as means to reach their objectives, while naïve users perceive these tools far from themselves. They think they are suitable for patients, or technology enthusiasts.
	Long-term usage	//	Rooksby et al. (2014) listed four different tracking styles for experienced users. We defined a new tracking style specifically addressed to naïve users.

Table 4. Comparison of our findings with previous research

These points highlight how naïve users have different needs, desires and attitudes from those of experienced users. They can be pushed in trying one PI tool by the suggestions of friends or significant others, the exposure to commercial ads, news or reviews, a casual glimpse to a device that attracts their attention. However, if these instruments will not be able to engage them immediately, quickly presenting how self-tracking could improve their daily life and wellness, and sustain their motivation, they will most be likely discharged and abandoned soon. In other words, they will never be able to discover the usefulness of collecting and exploring their personal data, because a premature abandonment will prevent them to develop goals and figure out how the information gathered could be integrated in their everyday activities. Experienced users know the potentialities of these kinds of technologies, but also their limits, and they strive to solve the problems they may encounter on their own; instead, naïve ones want all these issues to be managed by the tools themselves, and request proactivity, intelligence and immediacy from them.

6. IMPLICATIONS FOR DESIGN

As we showed, designing technologies to help naïve users gain self-awareness through self-monitoring is complex, since many practical issues affect their usage in everyday life. In this section, we want to move from the concrete findings yielded by our empirical work to the definition of a series of design strategies, intended as recommendations for the design of PI tools, in order to solve some of the most common barriers we encountered in their adoption. Empirical research can yield concrete results that can be applied to create design guidelines (Hekler et al., 2013), as done, for example, by Balaam et al. (2011), which defined a set of guidelines for motivating patients in stroke rehabilitation programs, starting from the findings collected during participatory design sessions. This type of work, thanks to its wide applicability to design, is an essential component of HCI research (Hekler et al., 2013).

We divided our design strategies in the four categories (tracking data, managing data, visualizing data, using data) derived from our diary study results (Section 4). For each category, we present some of the main issues found in our empirical work and the proposed design strategies to address them. Each design strategy is grounded in the findings gathered during the diary study and further supported by literature and examples, looking at different research fields, such as tangible interfaces, virtual environments and video games. We claim that applying insights coming from these fields could enhance PI tools making them more suitable to the users' needs. Some of these strategies explore paths not yet covered by any PI tool (1, 3, 4, 6). Others stress the importance of continuing along lines of research already explored within some applications: however they emphasize the need of improving or reinventing them, by suggesting, for example, to look beyond the most common gamification elements (7) already employed in some PI tools (e.g. Nike+), or by looking at user modeling techniques for solving privacy issues (2) and designing recommendations based on behavioral data (5). Finally, in proposing our strategies, we mainly focused on those issues that have not been addressed yet in other research applied to PI field. For example, significant steps ahead in regard to the integration of data were made by Bentley et al. (2013), who designed a system that correlates information from multiple sensor inputs and phone apps: then, we believe that it is not necessary to stress this point here.

To give them more concreteness, we start outlining a Persona, Marc, which will be then enacted in different scenarios to better describe the strategies suggested.

Scenario. *Marc is 33 years old, engaged, and lives alone in a small flat. He works as a librarian and has a quite sedentary lifestyle. He uses his desktop computer at home for internet browsing, playing and staying in touch with his friends. Even if he has a smartphone, he does not use all its features: he only has a few apps that he thinks are truly useful for him. Mark values wellness, happiness, friends, nature and travels. He feels he needs to do more physical activity and gain awareness of some aspects of his everyday life that sometimes prevent him from truly relaxing. These desires, however, have not been realized, since Marc does not believe that he has time, motivation and energy to explore himself deeper and make some kind of change in his habits.*

6.1 Tracking data

Issues. *Manually collecting data is cumbersome.* Although it is possible to imagine that in the future a variety of behavioral data will be tracked automatically thanks to the advancement of ubiquitous and wearable technologies, many aspects of people's behavior, such as emotions, will continue to rely on self-reporting, as they involve cognitive and interpretative components (e.g., emotional states still have a physiological part, which could be detected automatically by sensors, but also a cognitive part, which gives meaning to physical sensations and have to be interpreted and reported by the user). Thus, PI should find new ways to reduce the burden of self-monitoring.

Design strategy 1: *Remind and motivate the act of self-reporting*. Enhance user's engagement and motivation in reporting her data, by making the task of self-reporting "visible" and by fostering the development of an empathic relationship between the user and the PI tool. Use tangible interaction to involve people in tracking their personal data, as physical objects can remind them, by sheer presence, the act of self-reporting. On the other side, develop an emotional connection between user and tool, by exploiting the tendency of people to relate to computers in a social manner. By building systems able to trigger empathic reactions, users can be encouraged to report their data.

Rationale. This strategy aims at enhancing the involvement of users in tracking their own data. Tangible User Interfaces (TUIs) have emerged as a new type of user interface that leverages physical representation for linking the digital and physical worlds (Shaer and Hornecker, 2010). TUIs have proven to be more inviting than Graphical User Interfaces (GUIs) when a given task is not appealing enough on its own (Horn et al., 2010). It has been showed that TUIs provide a more engaging experience too, increasing the number of repeated activities accomplished by users (Xie et al., 2008). Using physical objects to track personal data can make this activity stimulating and enjoyable, since i) users are more physically involved, ii) the interaction can provide richer feedback, and iii) the experience is perceived as more real (Zuckerman et al., 2013). TUIs can materialize the task of reporting data in an object, reminding it to users simply through their physical presence. On the other side, research on virtual agents pinpoints how both the ability of expressing empathic emotions and the capability to trigger empathic reactions in users could bolster the interaction with computers, inspired by the way humans interact with each other (Ochs et al., 2012, Paiva et al., 2014). Studies conducted to assess how empathy affects people's attitudes towards robots found that participants who observed robots showing empathic behaviors perceived as closer the relationship with them (Cramer et al., 2010), and the robot itself as friendlier (Leite et al., 2013). Supporting the development of this emotional bond through design can encourage users to tell the tool how they feel and behave, by creating a trusted connection that can reflect the relation between therapist and patient.

Scenario. *A TUI positioned on Marc's desk reminds him, only with its presence, to track his emotional states every time he returns at home. The TUI shows him, through its colors, how much time it has passed since the last registration. Marc, then, by rotating the object, reports his data. This causes a change in the physical state of the object, which reflects his emotional state and shows an empathic response. For Marc, then, it is like telling something to a significant other, rather than to an inanimate artifact.*

6.2 Managing data

Issues. *Users perceive a lack of control on personal data.* Data privacy management is crucial in establishing users' trust toward PI tools. Losing control is the biggest fear among user. As long as conserving, sharing and deleting information are not considered under users' decision, it is hard to make them rely on the systems they are using. PI tools should give users the control on their personal information. Furthermore, people experience a sense of inadequacy in finding insights among the data collected. In clinical settings, interpretation is usually provided by the therapist/physician, while experienced users can rely on their experience in dealing with difficulties related on the management of their data. Naïve users, instead, are left alone in understanding their personal information.

Design strategy 2. *Give users the power over their data.* Give users the control on their data, making them perceive that they totally own what they have gathered. Leave them free to decide which data to store, where, for how long, to which purposes, whether to share and with whom.

Rationale. User control is the key to overcome privacy issues when we are dealing with many personal sensitive data, as happen in pervasive environments (Kay and Kummerfeld, 2006). Several studies (Knijnenburg and Kobsa, 2013; Wang and Kobsa, 2007) have showed that people's privacy preferences differ to some extent: there are some users that are more jealous of their personal information, while others are more willing to share them with other people. A solution in User Modeling (Brusilovsky, 1996) field is to make the user model *scrutable*, i.e. enabling users to inspect the model built on their data and to set their preferences about privacy management. Traditionally, *scrutable user model* is applied to e-learning systems in order to allow learner to inspect her goals (Bull and Kay, 2007). Allowing users to explore all the data the PI systems store about them, giving them the possibility to share what they like and “forget” what they want is important to improve trust towards these new kinds of technologies.

Scenario. *Marc has just discovered a unexpected correlation among his physical activity and his sleep quality. Although he wants to compare these data with those of his girlfriend, Elisa, he is not willing to expose them on Facebook. He decides to send these patterns to Elisa. After a quick look at what he has eaten the day before he decides to set the "memory” of the tool on this specific behavior to one day. He prefers his food habits to be stored only for a short time span.*

Design strategy 3. *Leave users free to help each other.* Allow users to become part of a small group of trusted friends, where they can share their personal information, compare what happened to them and find useful suggestions that can arise from the collective discussion.

Rationale. Naïve users require more intelligence in PI tools in managing their personal data. However, actual instruments seem not to be able either to highlight correlations relevant for the individual or to suggest which events are the causes of others. This strategy proposes to rely on the collective human intelligence to provide users with the insights they are searching for. In mutual-help groups, individuals that have a common problem meet regularly to support reciprocally. In these groups, all the members are considered peers and can give and receive help: friendship relationships among them can further relieve the stress factors that usually prevent the search for help (Humphreys et al., 1999). By having users in small groups with a common aim, in which each member can share her personal data with others, it is possible to compensate for the lack of intelligence of the current PI tools: individuals could then exchange their points of view, by suggesting each others correlations and causations based on their personal experiences. The creation of a trusted intimate space, fostered by the development of friendship relationships, could further invite users to expose themselves and their data, overcoming the widespread diffidence in making them public.

Scenario. *Looking at the trends of his monthly data, Marc has just discovered that when he goes for a run he often has trouble in falling asleep. However, by sharing these data in the group he belongs to, he becomes aware that another explanation is possible. Two friends of his left him a message, making him note that the problems related to his sleep are likely to connect with the type and the quantity of food he has eaten in those days, rather than with his physical activity. Thinking about it, Marc contacts them privately asking for further explanations. So, he convinces himself that the real problem relies in the excessive quantity of food that he is inclined to eat immediately after a run.*

6.3 Visualizing data

Issues. *Visualizations are neither meaningful nor tailored to the user needs.* People are not very familiar with visualizations of quantitative data that PI tools usually provide them. The data presentation should immediately engage users, giving sense to their self-monitoring activity. Furthermore, personalized goals, reports and suggestions on how they can address the issues found in the data collected would make the usage of PI tools more meaningful.

Design strategy 4. *Mirror the user.* Reflect user in an image that she can recognize, at the same time, as herself and as something else, as usually happens when a player identifies with her avatar and simultaneously feels a sense of attachment that leads her to improve it.

Rationale. Naïve users encounter difficulties in mirroring themselves in abstract representations of their data, finding them cold and aseptic. This strategy proposes to present the user's personal information in a way that can support her identification and promote the development of an emotional bond with the visualizations provided, in order to enhance user's self-reflection. Avatars give the user the possibility to reflect and identify herself in an alter ego. Avatar identification has shown to be positively correlated with enjoyment in games (Hefner et al., 2007) and satisfaction and retention in virtual worlds (Ducheneaut et al., 2009). Moreover, individuals tend to establish an emotional link with avatars, which resembles the level of intimacy one experiences when interacting with a close other (Ganesh et al., 2004). Data driven avatars, for example, could engage users in exploring their data supporting identification with them. Furthermore, they could also make data emotionally closer to users. Finally, showing both avatars based on current user behavioral data and ideal versions of them, representing the user's future and desired end states, could motivate her to put forth her best effort to achieve such ideal images, maintaining, for example, healthy lifestyles (Kim and Sundar, 2012).

Scenario. *Every evening, Marc visualizes his data. The system provides him with an impressionistic image of his self through an avatar, based on the data he collected during the day. Marc then explores in details the evolution of his alter ego in the last month, noting a positive trend in his physical activity. He wants to improve the characteristics of his avatar, and thus of himself, to reach the optimal state represented by the maximum possible grade of evolution of his digital representation*.

Design strategy 5. *Provide personalized reports, goals and suggestions exploiting narrative forms of presentation.* Provide users with personalized resumes of their daily data, objectives for improving their lifestyles, and hints related to how to achieve them, suggesting behaviors to follow and explaining the motivations of the recommendations provided. Exploit narrative mechanisms to present these information to the user, in order to enhance her engagement.

Rationale. Naïve users need to perceive the benefits that PI tools can provide them in their everyday lives. This strategy aims to support users in attaining their goals by providing personalized recommendations. PI tools could facilitate personal goal management, by employing a lifelong User Model (Kay and Kummerfeld, 2009), which has the potentiality of integrating data provided by different PI tools in a complex model of the user, that evolves on the basis of her behavior and can drive tailored suggestions for achieving personal goals even in the long term. In conjunction with techniques employed in *Recommender Systems* (Ricci et al., 2011), commonly used to reason on user data in order to suggest e.g. items to buy, or movies to see, a lifelong User Model could help users become aware of their faults and find a solution based on their past and current habits. Furthermore, as naïve users are bewildered by the overabundance of numeric information and require daily reports to understand the data they collected, this strategy suggests to present this information in a narrative form, and connect it with the recommendations provided. Narratives, in fact, not only played a fundamental role in all the cultures in preserving and transmitting information, but also provided those connections to the facts that can make them memorable (Austin, 2011). Data presentation techniques can then look at the ways through which the stories are narrated, in order to provide information (Gershon & Page, 2001).

Scenario. *Marc is in Vienna for a business trip. It is morning and he wants to go out for a jog. The PI system provides him with some short suggestions based on his user model. The system recommends Marc to run for 4 miles that day, as it knows his habits and his needs: this week Marc has not done a lot of physical activity and he needs to maintain at least last month's average in order to preserve his condition. The system explains the reasons behind this suggestion connecting it with the story of his past days. Then, it indicates a route that passes through three different gardens: the system knows, thanks to the user model, that Marc likes that setting.*

6.4 Using data

Issues. *Users need to relive the episodes connected to their data for gaining meaningful insights, while their motivation can rapidly decrease when they do not find immediate benefits in their everyday usage.* Naïve users showed that spatial cues are essential in the retrieval of memories that can enrich the meaning of the data collected: without remembering the episodes related to their data, users perceive them as scarcely useful for gaining insights. Moreover, on the one side PI tools show most of their benefits in the long term, when users are able to see the benefits of an increased self-knowledge. Nevertheless, many practical issues cause a rapid user disengagement. On the other side, PI tools can also be used for specific moments in time and for specific purposes. However, if users are not capable of understanding the value of their data even from their first self-tracking attempts they will never develop such purposes. PI tools should associate appealing meanings to the self-tracking activities and find ways to motivate users in the first phases of usage, preventing their immediate abandonment.

Design strategy 6. *Give the user the opportunity of reliving her data*. Provide users with contextual information that are able to trigger a reminiscence process, as memories are fundamental in adding meaning to the data collected. Rely especially on spatial cues for eliciting the user's memory, as space is a fundamental dimension strongly connected to the individual's past.

Rationale. Users need contextual cues for re-experiencing the data they gathered. This strategy proposes to sustain the episodic memory of the user, by anchoring her personal data to the spatial dimension. Episodic memory, in fact, is a personal registration of past events, which records information located in a time and in a space (Tulving, 1972). These types of memories are encoded according to a momentary context, mainly connected with what, where and when (Wheeler et al., 1997), which should be part of the reminiscence process. Following also the work of Li et al. (2012), which use contextual elements to support reflection on the factors that affect user's physical activity, this strategy instead stresses the importance of the spatial cues to trigger a reminiscence process that can enrich the meaning of the data collected. Spaces, in fact, are deeply connected with the people's emotional past experiences, as stressed by our users, and through them it is possible to favor the re-experience of the conditions in which particular data were tracked, adding more value to them. Research in location-based reminders also suggests the importance of space in triggering the remembering process (e.g. Beigl, 2000).

Scenario. *It is evening and Mark is looking at the data related to his mood. He notices, with no little surprise, that he had a peak of good mood three weeks before at seven o'clock. The system then shows him that in that moment he was running in the park, just before going to work, an unusual episode for Marc. Thanks to the cues that the system provides him, Marc remembers the air, the silence, the light and the sense of peace and wellness of that morning, and the fact that he intended to make more physical activity just at the beginning of the day. By becoming aware that he never followed this proposition, he starts to think that he should experience again those feelings in the next days.*

Design strategy 7. *Sustain user motivation in the whole user journey by leveraging both extrinsic and intrinsic motivations.* Use game elements to transform the core activity of self-monitoring in a playful experience. By providing different gamification strategies targeted to the different phases of the user journey, it is possible to trigger a rapid involvement, which could be helpful until users develop intrinsic motivations to support the use of PI tools.

Rationale. Naïve users showed that they search for an immediate gratification when using a PI tool and that they are not willing to wait for gaining the benefits promised. Using game elements in PI instruments can make the tracking experience more enjoyable and fun, changing the perception that users may have of it. Gamification (Deterding et al., 2011) has showed to be an effective strategy to enhance user motivation and participation in a variety of contexts (e.g. Barata et al, 2013, Cechanowicz et al., 2013). Relying on extrinsic motivation, usually fostered by this design technique (e.g., Mekler et al., 2013), can relief users from the burden of tracking data, allowing them to gain immediate rewards. This can overcome the "cold start" in using PI instruments, as long as users autonomously understand their importance and value. However, using only extrinsic rewards should be carefully considered, because of a potential detrimental effect on the users' intrinsic motivation, impoverishing the overall experience the system can provide (Deci et al., 1999). Nevertheless, how to employ game elements to elicit also intrinsic motivations in non-game contexts is started to be investigated: by leveraging the social aspects of games, such as cooperation among groups, and by providing rewards that incorporate different values, such as power and reputation, it is possible to satisfy the needs of competence, autonomy and relatedness that support the development of intrinsic motivations (Rapp, 2015). In this manner, it is possible to foresee different gamification strategies along the user journey, where in the first stages extrinsic rewards can quickly involve users, and in the later ones intrinsic motivations can support a long-standing engagement. Nevertheless, it is important to highlight how this engagement should be pursued in regard to the practice of self-tracking, rather than in relation with a single tool. Users can take advantage of a tool for a while and for a specific purpose, and then abandon it or change it with something else. However, if they will become aware of the value of collecting and using personal data, they could easily start again to self-track when a different need will arise. For this, the gamification strategies we pointed out here should be designed to be effective among different platforms and instruments, supporting the switching of PI tools, or the abandon and resume of the self-tracking activity.

Scenario. *Marc has just started to discover a new PI application. What he likes most is the fact that tracking, managing and visualizing his own data has become a sort of a game. At first he plays without caring much of the outcomes that this activity can have. As the time passes, he becomes aware of the benefits of knowing himself: the game is not played to obtain extrinsic prizes anymore but with the willingness and consciousness of doing something that produces valuable outcomes for him.*

7. CONCLUSION

Results of our study highlighted some criticalities experienced by naïve users in dealing with current PI tools. Naïve users showed poor compliance in self-monitoring. Without a strong initial motivation to investigate their own behaviors, self-tracking activities were dealt with scarce continuity, perseverance and accuracy. They most likely started with curiosity for a new class of tools that promised some kind of improvement in their daily lives. PI technologies were not able to keep these promises: many practical issues prevented users to become deeply engaged, leaving their potentialities hidden as well as their benefits in promoting self-awareness and self-knowledge. Naïve users wanted to minimize their efforts and maximize results: moved by unrealistic expectations, they became easily disappointed by what they felt to be cumbersome tasks, ambiguous representations and unintuitive interaction modalities.

We identified seven design strategies to overcome some of the issues highlighted throughout our empirical work. We are aware that these strategies do not address all the issues we found in our research. We only tried to take one step in this direction, proposing recommendations that aim at opening spaces of reflection that are deserving of further investigations. We hope that these strategies will benefit designers of PI systems, raising awareness of the intrinsic limits of the current PI tools, and showing them some directions to explore in subsequent research.

REFERENCES

Austin, M., 2011. Useful Fictions: Evolution, Anxiety, and the Origins of Literature. University of Nebraska Press, Lincoln.

Balaam, M., Rennick Egglestone, S., Fitzpatrick, G., Rodden, T., Hughes, A.-M., Nind, T., Axelrod, L., Harris, E., Ricketts, I., Mawson, S., Burridge, J., 2011. Motivating mobility: designing for lived motivation in stroke rehabilitation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). ACM, New York, 3073-3082.

Barata, G., Gama, S., Jorge, J., Gonçalves, D., 2013. Improving participation and learning with gamification. In Proceedings of the International Conference on Gameful Design, Research, and Applications (Gamification '13). ACM, New York, 10-17.

Beigl, M., 2000. MemoClip: A location-based remembrance appliance. Personal and Ubiquitous Computing, 4(4), 230-233.

Bentley F., Tollmar K., Stephenson P., Levy L., Jones B., Robertson S., Price E., Catrambone R., Wilson J., 2013. Health mashups: presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. ACM Transactions on Computer-Human Interaction (TOCHI), 20(5), 30:1-30:27.

Berners-Lee, T. James Hendler, and Lassila, O., 2001. Scientific American, 284(5), 34-43.

Bowen, G. A., 2008. Naturalistic inquiry and the saturation concept: A research note. Qualitative Research, 8(1), 137-152.

Braun, V., Clarke, V., 2006. Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.

Brusilovsky, P., 1996. Methods and techniques of adaptive hypermedia. User Model and User-Adapt. Interact. 6(2-3), 87-129.

Bull, S., Kay, J., 2007. Student Models that Invite the Learner In: The SMILI: Open Learner Modelling Framework. Int. Jour. of Art. Intel. in Educ., 17(2), 89-120.

Carter, S., Mankoff, J., 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). ACM, New York, 899-908.

Cechanowicz, J., Gutwin, C., Brownell, B., Goodfellow, L., 2013. Effects of Gamification on Participation and Data Quality in a Real-World Market Research Domain. In Proceedings of the International Conference on Gameful Design, Research, and Applications (Gamification '13). ACM, New York, 58-65.

Cheng, W. C., Golubchik, L., Kay, D. G., 2004. Total recall: Are privacy changes inevitable?. In Proceedings of CARPE 2004 Workshop. ACM, New York, 86-92.

Choe, E. K., Lee, N. B., Lee, B., Pratt, W., Kientz, J. A., 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). ACM, New York, 1143-1152.

Church, K., Smyth, B. 2009. Understanding the intent behind mobile information needs. In Proceedings of the international conference on Intelligent user interfaces (IUI '09). ACM, New York, 247-256.

Consolvo, S., McDonald, D. W., Toscos, T., Chen, M. Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J. A., 2008. 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). ACM, New York, 1797-1806.

Cramer, H., Goddijn, J., Wielinga, B., Evers, V., 2010. Effects of (in)accurate empathy and situational valence on attitudes towards robots. In Proceedings of the ACM/IEEE international conference on Human-robot interaction (HRI '10). IEEE, Piscataway, 141-142.

Crawford, K., Lingel, J., Karppi, T., 2015. Our Metrics, Ourselves: A Hundred Years of Self-Tracking from the Weight Scale to the Wrist Wearable Device. European Journal of Cultural Studies, 18(4-5), 479-496.

Czerwinski, M., Horvitz, E., Wilhite S., 2004. A diary study of task switching and interruptions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '04). ACM, New York, 175-182.

Deci, E. L., Koestner, R., Ryan, R. M., 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychological Bulletin, 125(6), 627-668.

Deci, E. L., Ryan, R. M. ,1985. Intrinsic motivation and self-determination in human behavior. Plenum, New York, NY.

Deterding, S., Dixon, D., Khaled, R., Nacke, L., 2011. From game design elements to gamefulness: Defining "Gamification". In Proceedings of the International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '11). ACM, New York, 9-15.

Dodge, M., Kitchin, R., 2007. Outlines of a World Coming into Existence: Pervasive Computing and the Ethics of Forgetting. Environment & Planning B: Planning & Design, 24, 431-445

Ducheneaut, N., Wen, M._H., Yee, N., Wadley, G. ,2009. Body and mind: a study of avatar personalization in three virtual worlds. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, 1151-1160

Elliot, A. J., Miltenberger, R. G., Kaster-Bundgaard J., Lumley, V., 1996. A national survey of assessment and therapy used by behavior therapists. Cognitive and Behavioral Practice, 3, 107-125.

Epstein, D. A., Lee, N. B., Bales, E., Fogarty, J., Munson, S. A., 2015. Wearables of 2025: Designing Personal Informatics at a Broader Audience. In proceedings of the Beyond Personal Informatics: Designing for Experiences with Data workshop at CHI 2015.

Foster, S. L., Laverty-Finch, C., Gizzo, D. P., Osantowski, J., 1999. Practical issues in self-observation. Psychological Assessment, 11, 426-438.

Fritz, T., Huang, E. M., Murphy, G. C., Zimmermann, T., 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). ACM, New York, 487-496.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J., 2009. UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, 1043-1052.

Ganesh, S., van Schie, H. T., de Lange, F. P., Thompson, E., Wigboldus, D. H., 2012. How the human brain goes virtual: Distinct cortical regions of the person-processing network are involved in self-identification with virtual agents. Cerebral Cortex, 22, 1577-1585.

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

Gershon, N., Page, W., 2001. What Storytelling Can Do for Information Visualization. Communications of the ACM, 44(8), 31-37.

Grinter, R., Eldridge, M., 2003. Wan2tlk?: everyday text messaging. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03). ACM, New York, 441-448.

Guarino, N. (Ed.), 1998. Formal Ontology in Information Systems. In Proceedings of FOIS'98. IOS Press, Amsterdam, 3-15.

Hammond, T. 2014. Wearables have a dirty little secret: 50% of users lose interest. Tech Republic Inc., February 2014. http://www.techrepublic.com/article/wearables-have-a-dirty-little-secret-most-people-lose-interest/ (accessed 1.03.16).

Hefner, D., Klimmt, C., Vorderer, P., 2007. Identification with the player character as determinant of video game enjoyment. Lecture Notes in Computer Science, 4740, 39-48.

Hekler, E.B., Klasnja, P., Froehlich, J.E., Buman, M.P. 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). ACM, New York, 3307-3316.

Herz, J. C., 2014. Wearables are totally failing the people who need them most. Wired, November 2014. http://www.wired.com/2014/11/where-fitness-trackers-fail/ (accessed 1.03.16).

Horn, M. S., Crouser, R. J., Bers, M. U., 2012. Tangible interaction and learning: the case for a hybrid approach. Personal Ubiquitous Comput. 16(4), 379-389.

Humphreys K., Mankowski E., Moos R. H., Finney J. W., 1999. Do enhanced friendship networks and active coping mediate the effect of self-help groups on substance abuse?. Annals of Behavioral Medicine, 21(1), 54–60.

Hunter, L. 2014 Are wearables over? Fast Company Inc., April 2014. http://www.fastcompany.com/3028879/most-innovative-companies/are-wearables-over (accessed 1.03.16).

IDTechEx, 2014. Wearable technology 2014-2024: Technologies, markets, forecasts e-textiles, wearable electronics, medical diagnostics, smart glasses, smart wristbands and more. July 2014. http://www.idtechex.com/research/reports/ wearable-technology-2014-2024-technologies\ -markets-forecasts-000379.asp (accessed 1.03.16).

Kay, M., Choe, E. K., Shepherd, J., Greenstein, B., Watson, N., Consolvo, S., Kientz, J. A., 2012. Lullaby: a capture & access system for understanding the sleep environment. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp '12). ACM, New York, 226-234.

Kay, J., Kummerfeld, B., 2006. Scrutability, user control and privacy for distributed personalization. In Workshop papers, CHI '06.

Kay, J., Kummerfeld, B., 2009. Lifelong User Modelling Goals, Issues and Challenges. In Proceedings of the Lifelong User Modelling workshop at UMAP 2009.

Khovanskaya, V., Baumer, E. P. S., Cosley, D., Voida, S., Gay, G. K., 2013. Everybody Knows What You're Doing: A Critical Design Approach to Personal Informatics. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). ACM, New York, 3403-3412.

Kim, Y., Sundar. S., 2012. Visualizing ideal self vs. actual self through avatars: Impact on preventive health outcomes. Comput. Hum. Behav., 28(4), 1356-1364.

Klasnja, P., Pratt, W., 2012. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. J. of Biomedical Informatics 45(1), 184-198.

Knijnenburg, B., Kobsa, A., 2013. Making Decisions about Privacy: Information Disclosure in Context-Aware Recommender Systems, ACM Trans. Interact. Intell. Syst. 3, 3, Article 20, 23 pages.

Korotitisch, W. J., Nelson-Gray, R. O., 1999. An overview of self-monitoring research in assessment and treatment. Psychological Assessment, 11, 415-425.

Ledger, D., McCaffrey, D., 2014. Inside wearables: How the science of human behavior change offers the secret to long-term engagement. Endeavour Partners, LLC 93(1), 36-45.

Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., Paiva, A., 2013. The Influence of Empathy in Human-Robot Relations. Int. J. Hum.-Comput. Stud. 71, 3, 250-260.

Li, I., Dey, A.K., Forlizzi, J., 2010. A Stage-Based Model of Personal Informatics Systems. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). ACM, New York, 557-566.

Li, I., Dey, A.K., Forlizzi, J., 2011. Understanding My Data, Myself: Supporting Self-Reflection with Ubicomp Technologies. In Proceedings of the international conference on Ubiquitous computing (UbiComp '11). ACM, New York, 405-414.

Li, I., Dey, A. K., Forlizzi, J., 2012. Using context to reveal factors that affect physical activity. ACM Trans. Comput.-Hum. Interact. 19, 1, Article 7 (May 2012), 21 pages.

Mann, S. 2004. Continuous Lifelong Capture of Personal Experiences with EyeTap. In Proceedings of the ACM workshop on Continuous archival and retrieval of personal experiences (CARPE'04). ACM, New York, 1-21.

Marcengo, A., Rapp, A. 2013. Visualization of Human Behavior Data: The Quantified Self. In Huang L. H., Huang, W. (Eds.) Innovative approaches of data visualization and visual analytics. IGI Global, Hershey, pp. 236-265.

Matthews, M., Doherty, G., 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). ACM, New York, 2947-2956.

Mekler, E. D., Brühlmann, F., Opwis, K., Tuch, A. N., 2013. Do points, levels and leaderboards harm intrinsic motivation? An empirical analysis of common gamification elements. In Proceedings of the International Conference on Gameful Design, Research, and Applications (Gamification '13). ACM, New York, 66-72.

Miltenberger, R. G., 2007. Behavior modification: Principles and procedures (4th ed.). Wadsworth, Belmont.

Nelson, R. O., Hayes, S. C., 1981. Theoretical explanations for reactivity in self-monitoring. Behavior Modification, 5(1), 3-14.

Ochs, M., Sadek, D., Pelachaud, C. A., 2012. formal model of emotions for an empathic rational dialog agent. Autonomous Agents and Multi-Agent Systems 24(3), 410-440.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Sobreperez, P., Woods, S., Zoll, C., Hall, L., 2004. Caring for Agents and Agents that Care: Building Empathic Relations with Synthetic Agents. In Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1 (AAMAS '04). IEEE, Washington, 194-201.

Rapp, A., 2015. Designing interactive systems through a game lens. An ethnographic approach. Computers in human behavior, doi:10.1016/j.chb.2015.02.048

Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (Eds.), 2011. Recommender Systems Handbook. Springer, Berlin.

Rooksby, J., Rost, M., Morrison, A., Chalmers M., 2014. Personal tracking as lived informatics. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14). ACM, New York, 1163-1172.

Rooksby, J., Rost, M., Morrison, A., Chalmers M., 2015. Pass the Ball: Enforced Turn-Taking in Activity Tracking. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, 2417-2426.

Salah, H., MacIntosh, E., Rajakulendran, N., 2014. Leveraging Canadian Innovation to Improve Health, MaRS Discovery District, Toronto, March 2014.

Shaer, O., Hornecker, E., 2010. Tangible user interfaces: past, present, and future directions. Found. Trends Hum.-Comput. Interact. 3(1-2), 1-137.

Sluijs, K. van der, Houben, G.-J., 2006. A generic component for exchanging user models between web-based systems. Int. J. Contin. Eng. Educ. Life-Long Learn. 16(1-2), 64-76.

Sohn, T., Li, K. A., Griswold, W. G., Hollan, J. D., 2008. A diary study of mobile information needs. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08). ACM, New York, 433-442.

Toscos, T., Faber, A., An, S., and Gandhi, M. P., 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). ACM, New York, 1873-1878.

Tudor-Locke, C., Bassett D. R., 2004. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med. 34(1), 1-8.

Tulving, E., 1972. Episodic and semantic memory. In Tulving, E., Donaldson, W. (Eds.). Organization of memory. Academic Press, New York, 382-403.

Wang, Y., Kobsa, A., 2007. Respecting Users' Individual Privacy Constraints in Web Personalization. In Proceedings of UM07. Springer-Verlag, Berlin, 157-166.

Wheeler, M. A., Stuss, D. T., Tulving, E., 1997. Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. Psychological Bulletin, 121(3), 331-54.

Xie, L., Antle, A., Motamedi, N. 2008. Are tangibles more fun? Comparing children's enjoyment and engagement using physical, graphical and tangible user interfaces. In Proceedings of the International conference on Tangible and embedded interaction (TEI '08). ACM, New York, 191-198.

Zuckerman, O., Gal-Oz, A., 2013. To TUI or not to TUI:. Evaluating Performance and preference in tangible vs. graphical user interfaces. Int. J. Hum.-Comput. Stud. 71(7-8), 803-820.