Leveraging User-made Predictions to Help Understand Personal Behavior Patterns

Miriam Greis
Tilman Dingler
Albrecht Schmidt
University of Stuttgart
Pfaffenwaldring 5a,
70569 Stuttgart
firstname.lastname@
uni-stuttgart.de

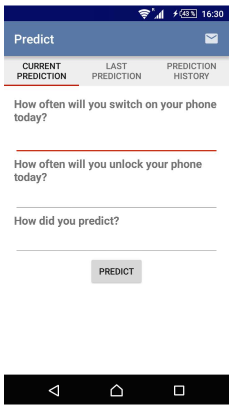
Chris Schmandt
MIT Media Lab
20 Ames St.,
Cambridge, MA 02139, USA
geek@media.mit.edu

MobileHCI '17, September 04-07, 2017, Vienna, Austria
ACM 978-1-4503-5075-4/17/09.
https://doi.org/10.1145/3098279.3122147

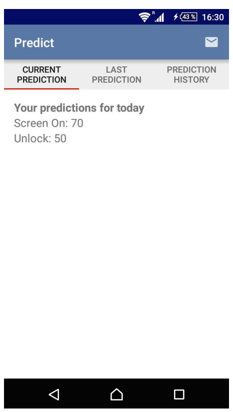
Abstract
People use more and more applications and devices that quantify daily behavior such as the step count or phone usage. Purely presenting the collected data does not necessarily support users in understanding their behavior. In recent research, concepts such as learning by reflection are proposed to foster behavior change based on personal data. In this paper, we introduce user-made predictions to help users understand personal behavior patterns. Therefore, we developed an Android application that tracks users' screen-on and unlock patterns on their phone. The application asks users to predict their daily behavior based on their former usage data. In a user study with 12 participants, we showed the feasibility of leveraging user-made predictions in a quantified self approach. By trying to improve their predictions over the course of the study, participants automatically discovered new insights into personal behavior patterns.

Author Keywords
User-made predictions; quantified self; personal behavior patterns; reflection

ACM Classification Keywords
H.5.m [Information Interfaces and Presentation (e.g. HCI)]: Miscellaneous



(a) Making a Prediction



(b) Prediction for the Current Day

Figure 1: Screenshots of the *Predict* application that show the current prediction screens.

Introduction

Smartphone applications and wearable devices (e.g., fitness bracelets) count how many steps we take, how many calories we burn, how many times we unlock our phones during the day, or how many words we read. Collecting all this data in the context of the quantified self movement (www.quantifiedself.com) leads to large personal datasets. Such datasets are often presented to users in order to motivate behavior change. This is semi-successful and prone to only bear short-term effects [10]. One reason for this is the abstract nature of the collected data: steps taken or calories burned do not necessarily relate to a metric users can grasp right away. Merely showing the collected data does barely inspire users to think about their behavior patterns related to those numbers. Nevertheless, understanding one's own behavior can be the first step towards behavior change. Therefore different strategies such as social competition, social support, rewards, goal setting, or learning by reflection are used to foster such behavior change.

In our work, we explore user-made predictions as a tool to improve the understanding of personal behavior patterns. We investigate whether asking users to predict their own behavior based on historical data makes them more aware of behavior patterns. For predicting, users need to think about the factors involved in a certain quantified result (i.e. correlations with real-life event such as social relationships). We decided to focus on mobile phone usage as smartphone addiction is becoming a more prominent issue. Additionally, mobile phone usage can be easily tracked and integrated in everyday life without participants having to wear additional sensors. We therefore built an Android application called *Predict* that tracks users' phone screen-on and unlock patterns. We conducted a user study with 12 participants who installed the application over the course of 2 weeks. Each day, we asked them to predict their phone usage behavior. Eventually, we collected their usage data logged by our application and asked them to fill in a questionnaire.

The contribution of this work is threefold: First, we assess the feasibility of leveraging user-made predictions in a quantified self approach. Second, we present a prototypical experiment investigating users' ability to predict their own phone usage behavior. Third, we provide quantitative and qualitative findings about users' prediction strategies and how user-made predictions can improve the understanding of personal behavior patterns.

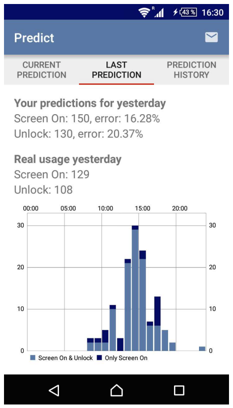
Related Work

In this section, we provide related work on the quantified self movement, its relation to persuasive computing and behavior change as well as related research about mobile phone usage.

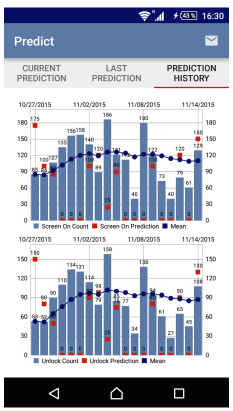
Quantified Self Movement

Technological advancements have given rise to the quantified-self movement. However, self-monitoring and self-tracking was had already been used for a while in behavioral psychology to allow subjects and clients to document their emotions and behavior [9]. Personal data collection with computers started in life logging research [6, 7]. Nowadays, sensors, websites, and mobile applications support users in collecting their personal data. Wearable devices from companies such as Fitbit (www.fitbit.com) or Jawbone (www.jawbone.com) track the personal step count, calorie expenditure, and other physiological data.

Fogg [5] had a huge influence on the research of the quantified self. Based on his vision of persuasive technology, many applications were developed. Most of them help to track health-related factors such as physical activity [4, 14], diet [18], sleep [8], or diabetes [16].



(a) History for the Last Day



(b) Complete History

Figure 2: Screenshots of the *Predict* application that show the history screens.

Besides the term quantified self, other terms such as personal informatics are in use. Li et al. [13] proposed a stage-based model for personal informatics systems, which consists of five stages: preparation, collection, integration, reflection, and action. Choe et al. [3] found that the collection and quantification of data is not enough, and that the goal of quantified-selfers is the reflection on the own behavior to possibly change the behavior. Self-reflection is therefore something that should be further examined in future research. Munson [17] proposed to decide between applications that want to change users' behavior and applications that only provide feedback for reflection but do not push the user to a certain goal.

Orji et al. [18] provide an overview of the most used theories from behavioral psychology and social sciences in research on persuasive technology. They further focus on three different categories for motivating users to change their behavior: social influence, learning and reflection, and game-like approaches. Rivera-Pelayo et al. [20] present a framework for the technical support of reflective learning in the context of the quantified self. They identify three support dimensions: tracking cues, triggering, and recalling and revisiting experiences. Tracking can either happen by self-reporting through the users or sensors that directly track the behavior. Tracked aspects can be any personal data of the user. A user can then be either actively or passively triggered by the system to reflect upon the collected data. Recalling and revisiting experiences is supported by contextualization, data fusion, data analysis, and visualization. As Baumer et al. [1] highlight, reflection also became a highly discussed topic in HCI during the last years.

Mobile Phone Usage

Understanding mobile phone usage is also a research topic gaining attention recently. Böhmer et al. [2] for example provide a detailed analysis on application usage on Android devices. They found that average application usage lasts less than one minute. Matching these findings, smartphone usage often consists of checking habits which may increase the overall phone usage [19]. Rising mobile phone usage also leads to problematic user behavior [11]. Lundquist et al. [15] found that college students had a more negative than positive view regarding the use and abuse of smartphones. To counter the overuse of smartphones, Lee et al. [12] propose design guidelines to support temporary non-use: fine-grained utility management, contextual interaction modes, self-regulation enhancement methods, and inconvenient interaction design. Two important aspects of self-regulation are self-monitoring and judgement of the personal behavior.

Several projects and applications already exist that show statistics about smartphone usage and try to reduce it. One very simple application is Checky (http://www.checkyapp.com/). Checky shows how often the users checked their phone. The application Moment (www.inthemoment.io/) tracks the total usage of IPhone users and allows to see how long family members used their phone. Menthal (https://menthal.org/) is a research project dedicated to track mobile phone usage behavior. Users can see how much they interacted with their phone on a day and can compare their usage to that of an average user.

Implications

Based on our review of related work, we identify reflection as an important aspect of the quantified self movement and as a design guideline to support smartphone non-use. New and different ways that enable people to reflect upon their data are necessary to generate a better understanding for personal behavior patterns. Therefore, we explore whether it is feasible to leverage user-made predictions in quantified

ID	Statement
S1	I liked to make prediction daily.
S2	I always looked forward to see how good my prediction was.
S3	I wanted to improve my prediction every day.
S4	I looked at the historic data before making a new prediction.
S5	I always used the same strategy for making my prediction.
S6	I think that my predictions improved over time.
S7	Making the predictions influenced my usage behavior.
S8	I tried to use my mobile device less.
S9	I will continue to make predictions with the app.
S10	I will continue to look at the historic data with the app.

Table 1: Statements presented to participants in the online questionnaire. Participants were asked to indicate their agreement with the give statements on a five-point Likert scale.

self approaches. We assume that making predictions based on historical data automatically animates people to think and reflect on their behavior. Additionally, it is a challenge that users are motivated to complete as best as possible. By trying to improve their predictions, we assume that users will automatically improve their understanding of personal behavior patterns.

Android Application

We developed an Android application that allowed users to predict their mobile phone screen-on and unlock patterns. Each day, when they for the first time used their mobile phone, users received a notification as a reminder to make a prediction, namely how often they would presumably turn on the screen and unlock their phone on this current day (see Figure 1a). As soon as they entered their prediction for the day, the application showed the predicted values (see Figure 1b). As support for predicting their usage, users had detailed statistics about their usage from the last day (see Figure 2a) and a summarized statistics with all their history of real usage and predictions (see Figure 2b). To extract the exact numbers of the usage behavior, the *Predict* application tracked the Android events SCREEN_ON, SCREEN_OFF, and USER_PRESENT and calculated how often the screen was turned on and how often the phone was unlocked. To preserve privacy, the application did not send any data to an external server unless participants pressed the button in the top right corner of the application. They then had the possibility to have a detailed look at their data before sending it. This ensured full control over personal data and made the data collection process transparent to the participants.

Method

We invited 12 prospective participants to be part of our study and asked them to install the application to use it on a daily basis for at least 14 days. After that time, we reminded them to send us their data and asked them to fill in an online questionnaire, which consisted of two parts: in the first required part, participants had to indicate their level of agreement on a five-point Likert scale with ten statements (see Table 1). In the second optional part, we asked them to give free feedback on different aspects such as whether the predictions influenced their usage behavior, and whether they learned something.

Participants

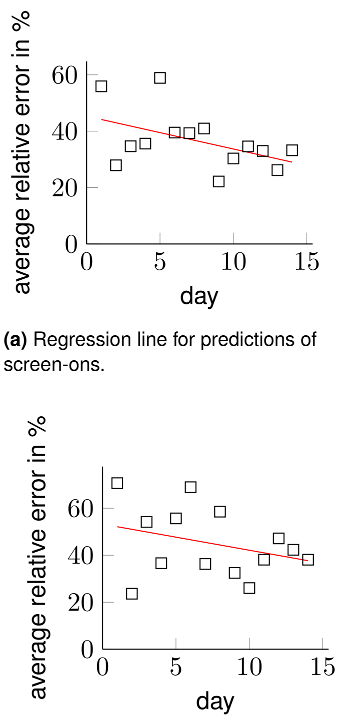
In total, we had 12 participants (10 male, 2 female) with an average age of 27 ($SD = 11.71$). Eight of them were students and the other four participants were employed for wages. All participants used the application for at least 14 full days, but most of them continued using the application voluntarily for at least a few more days.

Results

Over the course of 14 days, we registered a total of 9,317 unlock events, 6,576 additional screen-on events (no unlock was performed), and 336 predictions. In general, the number of screen-ons and unlocks varied highly across participants. Participants unlocked their phone between 2 and 264 times a day ($M = 55.79, SD = 41.74$) and additionally turned on their screen between 2 and 399 times a day ($M = 39.38, SD = 52.33$). In the following, we list the results for predictions made by participants regarding their screen-on and unlock prediction accuracies and how they behaved over time. For prediction accuracy, we calculated the relative error in percent.

Screen-on & Unlock Predictions

For the screen-on predictions, participants on average were off by 36.65% ($SD = 34.07\%$). For the unlock predictions participants on average were off by 44.94% ($SD =$



(a) Regression line for predictions of screen-ons.

(b) Regression line for predictions of unlocks.

Figure 3: Regression lines for the average relative error of screen-on and unlock predictions.

50.43%). Participants' predictions improved over time. On the first day, participants were off 55.92% ($SD = 34.78\%$) for screen-on predictions and 70.67% ($SD = 50.86\%$) for unlock predictions. Compared to the first day, the error rate on the last day of the study fell more than 20% for sreen-on predictions ($M = 33.23, SD = 31.29\%$) and more than 30% for unlock predictions ($M = 38.09\%, SD = 37.55\%$). Figure 3 illustrates the development of the average relative errors over time and the resulting regression lines. The regression lines indicate that although participants had more trouble in correctly predicting the number of unlocks, their predictions improved in the same manner as for the screen-on predictions.

Online Questionnaire
In the following, we present the results of the online questionnaire. We provide a detailed analysis of participants' Likert scale ratings (see Table 1 for the exact statements S1 to S10) and quotes. To better understand the ratings of the participants, we converted the ratings to numbers from 1 for "totally disagree" to 5 for "totally agree". Figure 4 summarizes all ratings of the participants.

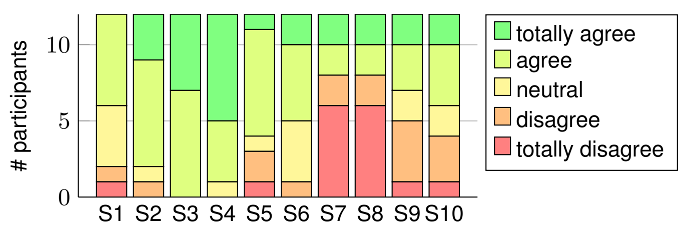


Figure 4: Results of the online survey showing the agreements of 12 participants on a five-point Likert scale (from totally agree to totally disagree) with the statements outlined in Table 1.

We asked participants what they liked about the application. Five participants liked the visualizations and diagrams, four participants found it easy to use and three participants liked the clean design of our application: "*I very much like the design, it is very minimal but super clear.*" (female, 30 years)

Half of the participants specified that they liked to make predictions daily, while most of the other participants were undecided ($S1, M = 3.25, SD = 0.97$). Only two participants did not like to make predictions. This came from different feelings that participants associated with the task of predicting their behavior. Participants who liked to predict their behavior found it interesting and challenging: "*At first I thought it would be annoying, but then I was surprised that it really was interesting and somehow challenging to better my predictions.*" (male, 22 years)

Participants that were neutral saw it more as an ordinary task and did not associate feelings with it: "*No special feeling. It was an ordinary task such as to set the alarm.*" (male, 28 years). The two participants who specified that they did not like it mainly got annoyed by the feeling that they did not get better in predicting their behavior: "*It got annoying over time. That I wasn't accurate at all didn't help.*" (male, 23 years)

Besides two participants, all participants stated that they looked forward to see how good their prediction was ($S2, M = 4.00, SD = 0.85$). Related to that, all participants wanted to improve their predictions from day to day ($S3, M = 4.42, SD = 0.51$), but only seven participants agreed with the statement that they think their predictions improved ($S6, M = 3.67, SD = 0.89$).

We further wanted to learn more about our participants' prediction strategies. Mostly, participants looked at the historic data before making their prediction ($S4, M = 4.5, SD = 0.67$). Eight participants agreed with the statement that they always used the same strategy to predict while four participants were neutral or disagreed ($S5, M = 3.41, SD = 1.16$). We asked participants to outline when they predicted and found that three of them predicted in the morning when getting up or unplugging their phone from the charger: "*Morning, when I unplugged the phone. Often just before I leave the house.*" (male, 63 years)

Five participants predicted directly after midnight while three more specified that they either predicted at midnight or in the morning depending on how long they stayed awake. Six explicitly referred to predicting when receiving the notification: "*I always predicted before I went to sleep somewhat after midnight. This was the first time the app remembered me to do so.*" (male, 25 years)

Additionally, we asked participants to provide us with more details on how they predicted. Participants specified that they started either with random predictions or used the data of the day before to make their prediction. One participant focused on the average after some days and specifically checked for the accuracy of the last day: "*I started with pretty random predictions for the first few days. Then as I saw a kind of average I focused on that. I guess after a week or so I thought about the day and what is going to happen. I also always checked for the last days accuracy because I thought of it as my average usage.*" (male, 21 years)

Most other participants tried to recognize patterns, use data from the same day the week before, or a day that should have a similar usag behavior than the day they try to predict on: "*I somehow tried to weigh different factors. First of all I made predictions using the last day. After I collected some data, I also used some patterns, I could recognize (e.g. specific day of the week or activities).*" (male, 20 years)

Other participants even figured out that more factors, such as their location and the weather, influence their usage behavior: "*At the beginning I used yesterday's historic data and prediction to make today's prediction. But that helps only part way because my phone use depends on many factors - day of week, whether in Boston or travel, whether I walk or bike to work, even whether it is raining (don't want the phone to get wet).*" (male, 63 years)

The influence of our application on the participants was perceived very differently ($S7, M = 2.33, SD = 1.67$). Eight participants disagreed that making the prediction had an influence on their usage behavior. One participants stated that he even forgot about the application until the next day: "*I didn't recognize any influence... after making the prediction I usually forgot about the app till the next day.*" (male, 23 years)

Four participants agreed that the application had an influence on their behavior. The same participants agreed that they tried to use their phone less ($S8, M = 2.33, SD = 1.67$): "*I tried to look less often at my phone. I have a notification light. So sometimes in the past I would still turn the screen on to double check if nothing is on. I wouldn't do that anymore because for 'predict' it would be counted. I also tried not to unlock it that often anymore or not unlock it randomly without having a real purpose apart from being bored. In general, I would say it supported me very much in being more aware of my phone usage as finally I had numbers that would back up how often I am using it.*" (female, 30 years)

We also wanted to know whether participants learned something by making the predictions and what they learned. One participant stated, that he did not learn anything. Another participants specified that he learned that his behavior was quite random. He was one of the participants frustrated by the application because his predictions were not accurate: "*My behaviour was quite random so I often predicted totally wrong.*" (male, 25 years)

Two other participants realized that they use their device more than they thought: "*That I looked at my device more than I thought I did.*" (female, 28 years)

Five participants explicitly stated that they learned and recognized patterns in their behavior such as where, when, and why they use their phone very often or not at all: "*That having my girlfriend around makes usage shrink. (what I see as a good thing)*" (male, 21 years) and "*I use the phone a lot less on the weekend, I hadn't quite realized how much less. How I travel makes a big difference on phone use. I don't turn it on as much as I thought I might.*" (male, 63 years)

We also asked participants whether they will continue making predictions with the application ($S9, M = 3.08, SD = 1.31$) and whether they will continue to look at the historic data of the application ($S9, M = 3.25, SD = 1.29$). From the six participants that agreed that they will continue to look at the historic data, five agreed that they will also continue to make predictions.

Discussion & Conclusion

We investigated the feasibility of leveraging user-made predictions in quantified self approaches and found that making predictions improves users' understanding of personal behavior patterns. As our quantitative and qualitative findings show, most participants enjoyed to make predictions and even saw them as challenge trying to improve their predictions daily. The analysis of the error rates supports participants feelings as they actually managed to improve their predictions over the course of the study. As assumed, most participants reflected on their personal behavior patterns to make better predictions gaining a deeper understanding of their usage. They learned about when and why they used their smartphones. Four participants even stated that they changed their usage behavior and tried to use their phone less, although behavior change was not one of our goals that we wanted to achieve with the application.

We see a big potential in leveraging user-made predictions for other types of personal data. Making daily predictions takes little time, but motivates people to reflect on usage patterns while trying to improve their predictions. We used screen-on and unlock events as preliminary examples, but are planning to extend our use case to app usages and to general information consumption. We imagine that game-like approaches or social competition can be combined with user-made predictions to increase users' motivation of making more accurate predictions which would eventually lead to more careful reflections on personal behavior patterns.

Acknowledgements

The authors would like to thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart

REFERENCES

1. E. P.S. Baumer, V. Khovanskaya, M. Matthews, L. Reynolds, V. Schwanda Sosik, and G. Gay. 2014. Reviewing reflection: on the use of reflection in interactive system design. In *Proc. DIS*. 93–102.

2. M. Böhmer, B. Hecht, J. Schöning, A. Krüger, and G. Bauer. 2011. Falling asleep with Angry Birds, Facebook and Kindle: a large scale study on mobile application usage. In *Proc. MobileHCI*. 47–56.

3. E. K. Choe, N. B. Lee, B. Lee, W. Pratt, and J. A. Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proc. CHI*. 1143–1152.
4. S. Consolvo, K. Everitt, I. Smith, and J. A. Landay. 2006. Design requirements for technologies that encourage physical activity. In *Proc. CHI*. 457–466.
5. B. J. Fogg. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002 (2002), 5.
6. J. Gemmell, G. Bell, and R. Lueder. 2006. MyLifeBits: a personal database for everything. *Commun. ACM* 49, 1 (2006), 88–95.
7. S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood. 2006. SenseCam: A retrospective memory aid. In *Proc. UbiComp*. 177–193.
8. M. Kay, E. K. Choe, J. Shepherd, B. Greenstein, N. Watson, S. Consolvo, and J. A. Kientz. 2012. Lullaby: a capture & access system for understanding the sleep environment. In *Proc. UbiComp*. 226–234.
9. J. Kopp. 1988. Self-monitoring: A literature review of research and practice. In *Social Work Research and Abstracts*, Vol. 24. 8–20.
10. D. Ledger and D. McCaffrey. 2014. Inside wearables: How the science of human behavior change offers the secret to long-term engagement. *Endeavour Partners* (2014).
11. U. Lee, J. Lee, M. Ko, C. Lee, Y. Kim, S. Yang, K. Yatani, G. Gweon, K.-M. Chung, and J. Song. 2014a. Hooked on smartphones: an exploratory study on smartphone overuse among college students. In *Proc. CHI*. 2327–2336.
12. U. Lee, S. Yang, M. Ko, and J. Lee. 2014b. Supporting Temporary Non-Use of Smartphones. In *CHI14 Workshop: Refusing, Limitng, Departing: Why We Should Study Technology Non-Use Workshop*.
13. I. Li, A. Dey, and J. Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proc. CHI*. ACM, 557–566.
14. J. J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H. B. Strub. 2006. Fish'n'Steps: Encouraging physical activity with an interactive computer game. In *Proc. UbiComp*. 261–278.
15. A. R. Lundquist, E. J. Lefebvre, and S. J. Garramone. 2014. Smartphones: Fulfilling the Need for Immediacy in Everyday Life, but at What Cost. *International Journal of Humanities and Social Science* 4, 2 (2014), 80–89.
16. L. Mamykina, E. Mynatt, P. Davidson, and D. Greenblatt. 2008. MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In *Proc. CHI*. 477–486.
17. S. Munson. 2012. Mindfulness, reflection, and persuasion in personal informatics. (2012).
18. R. Orji, J. Vassileva, and R. L. Mandryk. 2013. LunchTime: a slow-casual game for long-term dietary behavior change. *Personal and Ubiquitous Computing* 17, 6 (2013), 1211–1221.
19. A. Oulasvirta, T. Rattenbury, L. Ma, and E. Raita. 2012. Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing* 16, 1 (2012), 105–114.
20. V. Rivera-Pelayo, V. Zacharias, L. Müller, and S. Braun. 2012. Applying quantified self approaches to support reflective learning. In *Proc. LAK*. 111–114.