Challenges of Quantified-Self: Encouraging Self-Reported Data Logging During Recurrent Smartphone Usage

Aku Visuri[1], Niels van Berkel[2], Chu Luo[2], Jorge Goncalves[2], Denzil Ferreira[1], Vassilis Kostakos[2]
Center for Ubiquitous Computing[1], University of Oulu; [2]The University of Melbourne
[1](first.last)@oulu.fi; [2]n.vanberkel, chul3, jorge.goncalves, vassilis.kostakos@{student.}unimelb.edu.au

We argue that improved data entry can motivate Quantified-Self (QS) users to better engage with QS applications. To improve data entry, we investigate the notion of transforming active smartphone usage into data logging contributions through alert dialogs. We evaluate this assertion in a 4-week long deployment with 48 participants. We collect 17,906 data entries, where 68.3% of the entries are reported using the alert dialogs. We demonstrate that QS applications can benefit from alert dialogs: to increase data precision, frequency, and reduce the probability of forgetfulness in data logging. We investigate the impact of usage session type (*e.g.,* sessions with different goals or durations) and the assigned reminder delay on frequency of data contributions. We conclude with insights gathered from our investigation, and the implications they have on future designs.

Smartphones; Quantified-self; Interaction; alert dialogs

1. INTRODUCTION

It is commonly reported that users give up on quantified-self and fitness-tracking applications in a short timespan [7, 20]. Consumer surveys suggest a similar outlook, with a third of the surveyed 6,223 US adults reporting to have stopped using their activity tracker within just six months of ownership [11]. Previous work in attempting to increase the amount of self-reported data on smartphones include reminding the user via notifications [3] or text messages [15]. Both methods show increases in participation and interaction frequency. However, these methods can be time consuming, and when used in excess may frequently prompt the user to input data at inopportune moments.

The ease and timing of input for quantified-self data collection, is yet to be fully explored. Here, we investigate a way to transform idle time spent on smart devices into data contributions. Users of smartphones and wearable devices engage with their devices frequently during the day. The majority of smartphone use is unprompted and habitual [12], and the use is rarely planned in terms of continuing existing goals [19]. Offering the user a quick and efficient method for data contribution in tandem with these short bursts of micro-usage [5], or longer usages sessions, could lead into more effortless contributions.

To validate our claim, we conduct an investigation with a quantified-self application functioning as a proxy for data contributions. We use smartphone alert dialogs as a method to study the impact of continuous data logging on both the burden caused to the user, and the growth of data quantity gained

by enabling increased *in-situ* data contributions. We collect data over a period of four weeks.

The main contribution of our work is an understanding of which smartphone usage characteristics can be leveraged to increase daily data contributions. We present the insights gained during our deployment, and our proposals for future designs of interaction methods taking advantage of the knowledge of different types of device usage.

2. RELATED WORK

A common critique of activity and symptom trackers is that they are unable to sustain longitudinal user-tracking [9]. Recent research has shown that user motivation and application usage can be affected by enabling novel interactions with the user. Fritz *et al.* [6] studied the effects of long-term usage of quantified-self applications and report that applications that require more effort from the users are less likely to be successfully adopted and suffer from diminished application lifetimes [1]. The issue of applications being abandoned is non-trivial, as the concept of quantified-self relies on continuous tracking [17], and individuals are quick to return to their old habits once self-monitoring ceases [9].

A typical approach to increase self-monitoring is using reactive methods that remind the user to interact. Bentley *et al.* [3] used notifications to remind study participants to log eating and drinking habits. In comparison to a pilot study with no reminders they saw a rise from 20% daily logging ratio in study participants, to 63% of the participants remembering to log data daily when reminded with a notification. Tsai *et al.* [18] used

different reminder patterns via text messages and noted that while the prompts were answered, the increased data amounts were accompanied by higher feel of burden, and that their mobile interface was not appropriately designed for complex inputs. In [15], participants logged data with an average of 80% response rate when presented with two to five daily SMS prompts.

Alert dialogs (per Android terminology), or more commonly used term popups, are an UI element sparingly used due to their interrupting nature. Alert dialogs are occasionally used in tandem with the Experience Sampling Method [10]. van Berkel *et al.* [19] used alert dialogs to log ESM responses from their participants, and report average response times of 2.70 seconds and response rate of 83.78%. Akhawe *et al.* [2] studied browser warning dialogs and the theory of warning fatigue and conclude that the UX of a warning dialog can have a tremendous impact on the user behaviour and experience.

Due to their small size, alert dialogs do not necessarily involve a heavy cognitive load which, according to Leiva *et al.* [13], is one of the two major issues with mobile information presentation methods. Additionally, the use of alert dialogs as input mechanisms on smartphones, rather than information presentation methods, is not comprehensively studied.

The hypothesis for our investigation is the following: by detecting naturally appearing breakpoints (*e.g.* micro- breaks), or periods of smartphone use where the user is not actively performing a task, *i.e.* juggling applications while bored [14] and aimlessly [4], **we can transform active device use into data contributions**.

Evidently, the length of such interactions is kept to a minimum, to not substantially prolong shorter usage sessions, and to not needlessly burden the user. The design of interaction methods leveraging these brief usage sessions should be quick (C1), unobtrusive (C2), easily ignorable (C3) by the user when required, and should not involve delays (C4) [8]. Based on inherent properties, alert dialogs can potentially fill the conditions described in Table 1.

3. METHODOLOGY

We developed an application called LifeTracker that enables the user to contribute data using alert dialogs as an input method. We then recruited 48 participants to use the application in a 4-week long user study. In our work, we focus on the nature of alert dialogs as an input mechanism for an application of this type, and do not analyse other aspects of the application, such as data visualisations or potential target audiences.

3.1 LifeTracker Application

To investigate the use of alert dialogs as an input method to unobtrusively increase data quantity in quantified-self applications, we developed an Android application called LifeTracker.

Table 1: Design requirements for leveraging usage sessions for data contributions.

Condition	Description
C1	*Alert dialogs can be interacted with very briefly.*
C2	*Alert dialogs can be presented to the user at appropriate times (with the use of interruptibility detection methods).*
C3	*Alert dialogs are easy to dismiss.*
C4	*Alert dialogs can be quickly generated.*

The application has three different tracking schemas (*e.g.*, mood tracking). Each of these schemas consists of two variables that can be inputted repeatedly (once per hour) and 7-9 variables that are logged once per day. *E.g.*, the exercise scheme asked participants on their number of daily stretching sessions (daily) as well as their experienced level of dehydration (hourly), among other variables.

Data can be inputted to the application via the alert dialogs (*Figure 1*), or directly from the application. The application and the alert dialogs include three input modalities (three-tier scale, multiple choice, and a range of values). The modalities were designed to be quick and convenient to interact with to ensure swift and effortless data logging (as per C1).

When prompted via an alert dialog, user can opt to input the data from the dialog, and interact with the dialog by either accepting or dismissing the dialog, or launching the application. If there is no user interaction, each alert dialog is automatically removed from the view after one minute.

The alert dialogs can be dismissed with a single interaction (fulfilling C3), and the interactions are designed to also enable data logging with single interactions (fulfilling C1). Alert dialogs are never generated when an attention consuming task, such as a phone call, is taking place (partly fulfilling C2), and do not require loading times or any background computations (C4).

The application also sends a daily reminder notification after 6pm if the user is missing any daily inputs, in order to compare to the method originally presented in [3] within our own study setting. Clicking the notification launches the application's main interface.

While the user is actively (screen is in unlocked state) using his smartphone, he can be prompted to

input **missing** data (data not yet inputted for particular day or hour) via an alert dialog in two ways. First, every time user unlocks his smartphone, he has a probability P of receiving an alert dialog (Figure 1 top-right). Secondly, after five minutes of uninterrupted (device remains unlocked) device use, he has probability P of receiving an alert dialog. Probability P begins at 50% for all users, and shifts within 5 – 95% according to the user's interactions with previous dialogs (*i.e.* frequent dismissal reduces P, and acceptance increases P).

3.2 Study Setting

We recruited 48 participants (36 males, 12 females, aged 21-53 years old, M = 28.04, SD = 6.71) using mailing lists at our University on a first-come first-served basis. The participants had varying academic background, ranging from Humanities and Law to Engineering and Natural Sciences.

In the study intake session, we explain how the application works, and how the data is tracked and stored remotely. The anonymity of each participant and their data was explained, and each of the participants was asked to read and sign a consent form. We requested the participants to use the application as they felt beneficial to their personal needs and that were not required to log data or use the application when it was inconvenient to them. We explained that the alert dialog generation in the application changes based on the way the participant interacts with the previously generated ones.

Figure 1: View from LifeTracker application's alert dialog requiring data input.

Following the 4-week long study period, we invited participants to complete a post-study questionnaire regarding their experience with the application.

4. RESULTS

In our four-week long investigation, we collect a total of 17,906 data entries from 48 users. A majority (68.3%, N = 12,222) of the data was logged using alert dialogs, and the remainder of the data (31.7%, N = 5,684) was logged directly from the application's main interface.

Users responded to an average of 13.63 alert dialogs per day (N = 19,271). Responses contain 15,424 (80.0%) interactions labelled as *accept*, 3652 (18.9%) interactions labelled as *dismiss*, and 195 (1.0%) events where the application was launched directly from the dialog. Accepting or dismissing a dialog does not require data contributions, which explains the discrepancy between the number of data logged from dialogs (N = 12,222) and the number of accepted dialogs (N = 15424). It takes user average of 5.08 seconds to complete an interaction with an alert dialog (excluding dialogs that were dismissed without contributions).

As indicated by these numbers, the users generally perceived the dialogs favourably (80% accepted). Part of this behaviour can be attributed to the fact that users were aware that logging data will make subsequent alert dialogs deferred to a later time. However, as described later in the survey results, the users did not experience that this behaviour was due to the potential future burden.

4.1 Combined Notification and Alert Dialogs

2,655 entries were inputted from the main interface after 6pm, without a prior dialog. These interactions were user initiated application launches or result of the reminder notification. During evening hours, a similar amount of data was collected from the dialogs (N = 3,481). During earlier hours majority of the data was logged using alert dialogs (N = 8,741 for dialogs, N = 3,029 from application). As the two methods are not directly exclusive of each other, we can observe increase of 131.1% in data contributions during the hours when user received both alert dialogs and a reminder notification over just using a notification reminder. This result is an additional improvement to the use of reminder notifications, as presented in [3].

4.2 Device Usage

The users logged 395,204 screen events which are aggregated into 49,296 usage sessions. After removing outliers, the remaining sessions have an average length of 2 minutes 15 seconds and median length of 44.6 seconds. Users logged on average 53.7 daily usage sessions which is in line with previous work [21].

We use a threshold of 45 seconds between sessions, as presented in [19], to distinguish between usage sessions in which a user begins a new task and sessions in which the user continues her previous tasks. Our classification shows 35,000

(71.0%) new sessions and 14,296 (29.0%) continuing sessions.

5,625 of all sessions (11.41%) include data entries, and using McNemar's chi-squared test (x^2 = 38628.05, φ = 0.035, p < .05) we verify that sessions that begin new tasks (4,359 data entries, 8.8% of all, 12,5% of sessions with new task) are more likely to result in data contributions than sessions continuing previous tasks (1,367 entries, 2.8% of all, 9.5% of continuing sessions).

Based on the distribution of session lengths and frequency of data contributions in said sessions, longer usage sessions are more likely to lead into data contributions (0.39 > 0.12 contributions per session, p < .05). Sessions longer than five minutes (the threshold for the delayed alert dialog) lead to average of 0.39 data contributions per session.

Users are more likely to be disrupted when they are working on a task (continuous session) rather than when they begin a new task. The delayed alert dialogs (N = 644, 5 minutes after unlocking) prompted more favourable interactions. The differences are statistically significant (x^2 = 45.117, Cramer's V = 0.068, p < 0.05). Delayed dialogs prompted 46.7% less dismissals and led to 186% more direct application launches. Clearly when users have more time available, they are more willing to contribute.

Table 2: Summary of data contributions.

Data contributions (N = 17,906)		
Alert dialog	*68,3%*	
Application	*31,7%*	
	Before 6PM (N = 11,770)	**After 6PM (N = 6,136)**
Alert dialog	*74,3%*	*56,8%*
Application	*25,7%*	*43,2%*

Table 3: Summary of the distribution of accepted alert dialogs.

Accepted dialogs (N = 19,271)	
Overall	*80,0%*
Delayed	*89,4%*
'Instant'	*79,6%*

Table 4. Summary of the usage sessions.

Sessions (N = 49,296)	New task	Continuing
Overall (N = 49,296)	*71,0%*	*29,0%*
Including contributions (N = 5,625)	*12,5%*	*9,5%*

4.3 Post-Study Survey

53.3% of the responses preferred alert dialogs for data logging, 22.2% preferred to use the application, and 24.4% were undecided. From the positive responses, we classified the following four repeatedly occurring themes:

Remembering data input: *"Without the alert dialogs, I don't think I would have remembered to answer anything."*

Instantaneous nature: *"I prefer dialogs because they are throughout the day, and my mood can change several times during day."*

Effortlessness: *"Alert dialogs, because it took less effort.",* and *"Alert dialogs were easier."*

Distributed logging: *"The way it is submitted in bits which feels like it takes less time compared to submitting all of the data at one point."*

Of these two, the interrupting nature of alert dialogs is likely due to the nature of interruptive interaction methods, such as notifications as alert dialogs, which can be seen as interruptive if the user believes that the content or the source of the interruption is not beneficial to his or her [16].

5. DISCUSSION

Development of efficient interaction methods for applications to continuously collect self-reported data, and the related issue of continuous data collection is one of the core problems of longitudinal use of quantified-self applications. We investigate if the behaviour of smartphone users to interact with their devices frequently [21] could be put into use by transforming part of this device use into data contributions. We design a method to prompt the user to log data via alert dialogs during active device use. Similar, but delayed reminder methods are used by researchers in the form of notifications [3] and text messages [15].

We investigate the use alert dialogs to both enable and encourage users to contribute data more frequently in a four-week study with 48 users. Over this time, users responded to 644 delayed dialogs and 18,627 dialogs generated directly after unlocking their device. majority of the data - 68.3%, N=17,906 - is logged via alert dialogs, and the remainder directly from the application.

5.1 Reflecting on Previous Work

These results are not directly comparable to previous work, *e.g.*, [3], but some comparisons can still be drawn. In [18], participants interacted with the application on average 6 times per day, which is 56.3% of the *accepted* alert dialogs (M = 10.64 per day) generated by LifeTracker. In addition to this, our users log 31.7% of all data via the application indicating frequent application launches (*Table 2*). Our reminder dialogs further increase the amount of logged data when they were presented in tandem with a reminder notification.

When asked for reasons for selecting the input method, participants agreed on four different aspects: 1) **improved recall**, 2) **in-situ** reports, 3) **low effort**, and 4) more **evenly distributed** self-reports throughout the day. Improved recall is an inherent aspect of all reminders - be it an alert dialog, notification, or a text message. Situatedness and distribution are longitudinal in nature. As many measured aspects of our lives can change throughout the day (*e.g.*, mood) it is more reliable to request the user to report their current experience instead of historical information. This allows the application to generate more fine-grained insights on the user's own data.

Overall, most the users perceived dialogs as helpful. To understand their underlying reasons for interactions, we analyse their device usage during the study and how this use is reflected on their interactions.

5.2 Effect of Usage Sessions Type on Data Contributions and Interactions

We aggregate the screen events from users to 49,296 usages sessions. Of the average 53.7 daily sessions, more contributions are during sessions beginning a new task (*Table 4*). This behaviour can be attributed to the fact that a large portion of device use is not goal-orientated or is recreational in nature [4]. The sessions that start new tasks are over three times more likely to lead to data contributions (*Table 4*), which should be taken strongly into consideration as a method to prevent unnecessary disruptions and increase the number of data contributions.

Another factor impacting the probability of a data contribution is the session duration, as brief sessions are less likely to lead to data contributions or *accepted* alert dialogs. In longer sessions, users are more willing to contribute and the increase of session length correlates with the frequency of data contributions. This behaviour of increased contributions during longer usage sessions can apply to other interaction methods as well. When users engage with their devices for longer periods of time, the use is likely to be less time constraining, thus the prompt is less distracting in the beginning of the use. Additionally, during longer sessions the user is more willing to contribute (0.39 per session, 0.31, 0.12, for sessions lasting t = 5 minutes, t > 40s, t < 40 s, respectively). Hence, systematically predicting session lengths can improve data contributions.

While users should be encouraged to contribute more data, constantly reminding them causes a significant burden which can elicit negative attitudes towards the reminder method and source. Our analysis showcases that users are *more willing to contribute* during sessions where they **are not**

continuing previous goals or tasks, and during **longer usage sessions** which provide user with more available time or are more likely to be associated with recreational device use. We also show that **delayed reminders** during **active device use** are more favourably received.

5.4 Limitations

Our study period is relatively short in terms of quantified-self applications and their long-term benefits and usability to the users. Secondly, being a part of a study setting could cause the participants to change their device use from how they would normally use them.

6. CONCLUSION AND FUTURE WORK

We set out to investigate ways to improve quantified-self and other applications relying on self-reports by transforming recurrent smartphone usage into data contributions. We conducted a four-week long study with 48 users and analysed how they used their smartphones during this period. We investigate how frequently, during what type of device use, and via what interaction medium the data was logged.

Our results show that the amount of self-reports can be increased with the use of reminders containing user interfaces that enable data contribution, that are presented to the user during active device use. Reminders presented during longer usage sessions lead to more data contributions. We propose the use of reminders during usage sessions where user is starting a new task, and in the beginning or during longer sessions.

Constantly requiring the user's attention can still cause a burden, especially in long-term use, as not all our users reacted favourably to the continuous alert dialogs. More covert interaction methods should also be considered, that leverage the same concept of enabling repeated entries, and reduce the cognitive and disrupting load of an on-screen interface.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] A. Ahtinen, E. Mattila, A. Vaatanen, L. Hynninen, J. Salminen, E. Koskinen and K. Laine. Year. User experiences of mobile wellness applications in health promotion: User study of Wellness Diary, Mobile Coach and SelfRelax. In *International ICST Conference on Pervasive Computing Technologies for Healthcare*, IEEE, 1-8. DOI: https://doi.org/10.4108/ICST.PERVASIVEHEALTH 2009.6007

[2] Devdatta Akhawe and Adrienne Porter Felt. Year. Alice in warningland: A large-scale field study of browser security warning effectiveness. In *Presented as part of the 22nd USENIX Security Symposium (USENIX Security 13)*, Proceedings of the 22nd USENIX Security Symposium, 257-272.

[3] Frank Bentley and Konrad Tollmar. Year. The Power of Mobile Notifications to Increase Wellbeing Logging Behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1095-1098. DOI: https://doi.org/10.1145/2470654.2466140

[4] Barry Brown, Moira McGregor and Donald McMillan. Year. 100 Days of iPhone Use: Understanding the Details of Mobile Device Use. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services*, ACM, 223-232. DOI: https://doi.org/10.1145/2628363.2628377

[5] Denzil Ferreira, Jorge Goncalves, Vassilis Kostakos, Louise Barkhuus and Anind K Dey. Year. Contextual Experience Sampling of Mobile Application Micro-Usage. In *International Conference on Human-Computer Interaction with Mobile Devices and Services*, ACM, 91-100. DOI: https://doi.org/10.1145/2628363.2628367

[6] Thomas Fritz, Elaine M Huang, Gail C Murphy and Thomas Zimmermann. Year. 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*, ACM, 487-496. DOI: https://doi.org/10.1145/2556288.2557383

[7] Rúben Gouveia, Evangelos Karapanos and Marc Hassenzahl. Year. How Do We Engage with Activity Trackers?: A Longitudinal Study of Habito. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 1305-1316. DOI: https://doi.org/10.1145/2750858.2804290

[8] Damian Hoy, Lyn March, Peter Brooks, Fiona Blyth, Anthony Woolf, Christopher Bain, Gail Williams, Emma Smith, Theo Vos, Jan Barendregt, Chris Murray, Roy Burstein and Rachelle Buchbinder. 2014. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Annals of the Rheumatic Diseases*. annrheumdis-2013-204428. DOI: https://doi.org/10.1136/annrheumdis-2013-204428

[9] Evangelos Karapanos. 2015. Sustaining User Engagement with Behavior-change Tools. *interactions*, 22 (4). 48-52. DOI: https://doi.org/10.1145/2775388

[10] Reed Larson and Mihaly Csikszentmihalyi. 2014. The Experience Sampling Method. In *Flow and the Foundations of Positive Psychology*, Springer Netherlands, 21-34.

[11] Dan Ledger and Daniel McCaffrey. 2014. Inside Wearables *How the Science of Human Behavior Change Offers the Secret to Long-Term Engagement*, Endeavour Partners, 17.

[12] Yu-Kang Lee, Chun-Tuan Chang, You Lin and Zhao-Hong Cheng. 2014. The dark side of smartphone usage: Psychological traits, compulsive behavior and technostress. *Computers in Human Behavior*, 31. 373-383. DOI: https://doi.org/10.1016/j.chb.2013.10.047

[13] Luis Leiva, Matthias Böhmer, Sven Gehring and Antonio Krüger. Year. Back to the app: the costs of mobile application interruptions. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services - MobileHCI '12*, New York, New York, USA, 291. DOI: https://doi.org/10.1145/2371574.2371617

[14] Aleksandar Matic, Martin Pielot and Nuria Oliver. Year. Boredom-computer Interaction: Boredom Proneness and the Use of Smartphone. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM, 837-841. DOI: https://doi.org/10.1145/2750858.2807530

[15] Kevin Patrick, Fred Raab, Marc A Adams, Lindsay Dillon, Marian Zabinski, Cheryl L Rock, William G Griswold and Gregory J Norman. 2009. A text message-based intervention for weight loss: randomized controlled trial. *J Med Internet Res*, 11 (1). e1. DOI: https://doi.org/10.2196/jmir.1100

[16] Alireza Sahami Shirazi, Niels Henze, Tilman Dingler, Martin Pielot, Dominik Weber and Albrecht Schmidt. Year. Large-scale Assessment of Mobile Notifications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 3055-3064. DOI: https://doi.org/10.1145/2556288.2557189

[17] Melanie Swan. 2013. The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data*, 1 (2). 85-99. DOI: https://doi.org/10.1089/big.2012.0002

[18] Christopher C Tsai, Gunny Lee, Fred Raab, Gregory J Norman, Timothy Sohn, William G Griswold and Kevin Patrick. 2007. Usability and Feasibility of PmEB: A Mobile Phone Application for Monitoring Real Time Caloric Balance. *Mob. Netw. Appl.*, 12 (2-3). 173-184. DOI: https://doi.org/10.1007/s11036-007-0014-4

[19] Niels van Berkel, Chu Luo, Theodoros Anagnostopoulos, Denzil Ferreira, Jorge Goncalves, Simo Hosio and Vassilis Kostakos. Year. A Systematic Assessment of Smartphone Usage Gaps. In *Conference on Human Factors in Computing Systems*, ACM, 4711-4721. DOI: https://doi.org/10.1145/2858036.2858348

[20] N. van Berkel, C. Luo, D. Ferreira, J. Goncalves and V. Kostakos. 2015. The Curse of Quantified-Self: An Endless Quest for Answers. In *International Joint Conference on Pervasive and Ubiquitous Computing (Adjunct)*, 973-978. DOI: https://doi.org/10.1145/2800835.2800946

[21] Aku Visuri, Zhanna Sarsenbayeva, Niels van Berkel, Jorge Goncalves, Reza Rawassizadeh, Vassilis Kostakos and Denzil Ferreira. 2017. Quantifying Smartwatch Use. In *Conference on Human Factors in Computing Systems*, ACM.