Participants' Experience and Adherence in Repeated Measurement Studies Among Office-Based Workers

Junoš Lukan
Jožef Stefan Institute
Department of Intelligent Systems
Ljubljana, Slovenia
Jožef Stefan International
Postgraduate School
Ljubljana, Slovenia
junos.lukan@ijs.si

Larissa Bolliger
Ghent University
Department of Public Health and
Primary Care
Ghent, Belgium
larissa.bolliger@ugent.be

Els Clays
Ghent University
Department of Public Health and
Primary Care
Ghent, Belgium
els.clays@ugent.be

Oscar Mayora
Fondazione Bruno Kessler Research
Institute
Trento, Italy
omayora@fbk.eu

Venet Osmani
Fondazione Bruno Kessler Research
Institute
Trento, Italy
vosmani@fbk.eu

Mitja Luštrek
Jožef Stefan Institute
Department of Intelligent Systems
Ljubljana, Slovenia
Jožef Stefan International
Postgraduate School
Ljubljana, Slovenia
mitja.lustrek@ijs.si

ABSTRACT

While diary studies, especially when applying ecological momentary assessment (EMA), are a great way to capture self perceptions to later use as labels for other data, they can be a burden for study participants. To increase their adherence to the study design, it is important to tailor it to their needs and take their feedback into account. This paper reports on a data collection process in a study focused on occupational stress. The data collection is briefly described and the participants' responses are analysed in terms of adherence. Participants' feedback was collected at the end of the study and its main themes are summarized. These experiences are compared to the ones in another study focusing on stress and burnout, with a very similar methodological design. Some general conclusions are drawn from both with suggestions on how to best carry out an EMA study.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**; **Ubiquitous and mobile computing design and evaluation methods**; • **Applied computing** → Psychology.

KEYWORDS

occupational stress; data collection; ecological momentary assessment; adherence

UbiComp-ISWC '21 Adjunct, September 21–26, 2021, Virtual, USA

ACM ISBN 978-1-4503-8461-2/21/09…$15.00
https://doi.org/10.1145/3460418.3479367

ACM Reference Format:
Junoš Lukan, Larissa Bolliger, Els Clays, Oscar Mayora, Venet Osmani, and Mitja Luštrek. 2021. Participants' Experience and Adherence in Repeated Measurement Studies Among Office-Based Workers. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers (UbiComp-ISWC '21 Adjunct), September 21–26, 2021, Virtual, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3460418.3479367

1 INTRODUCTION

Occupational stress is a common occurrence in Europe and elsewhere and it has a proven health impact. For example, tight deadlines are experienced by one third of workers in the EU and one in five needs to work in their free time to meet these high demands [3].

We carried out a study to explore the relationship between work environment risk factors and stress outcomes in academic settings in a project called *Stress at Work (STRAW)*. We have previously published a protocol of the study [2]. In this paper, we report on the process of data collection and participants' adherence to the protocol as well as their feedback. We compare these to findings from a similar project called Turnout Burnout with the aim to draw more general conclusions and provide suggestions on how to do a study like these two well.

2 PROTOCOL DESCRIPTION

In the STRAW project, we collected three main types of data using different tools. The participants were asked to wear the Empatica E4 wristband measuring physiological parameters and fill in questionnaires for 15 working days. During these days, smartphone data from various phone sensors as well as phone usage data were collected using an Android application based on the AWARE framework [5]. No data was collected on weekends.

2.1 Ecological Momentary Assessment (EMA)

Participants responded to questionnaires every 90 minutes during their workday and one additional questionnaire in the evening. These ecological momentary assessments (EMAs) included items from established psychological and medical scales. They were selected to capture important work environment risk factors (based on the 6th European Working Conditions Survey [3]) and possible stress outcomes (described by Ice and James [7]). The EMAs offered during the workday consisted of questions related to work environment, while an evening EMA also asked about detachment from work and a whole working day impression. Each EMA contained two items per scale, selected at random each time to increase engagement. The daytime EMAs consisted of around 20 items, while the evening EMA of around 40, so that they could be answered in about 2 and 5 minutes time, respectively.

The EMAs' schedule was flexible and partly adapted to individual participants' needs. The workday EMAs were triggered randomly after participants marked the start of their work and then approximately every 90 minutes with no two EMA sessions closer than 30 minutes. This continued until participants marked the end of their workday. Finally, the evening EMA appeared at the pre-set time, which was recommended to be well after work hours but before bed time (e.g., at 20:00).

The participants were reminded of EMAs with a notification that repeated after 15 minutes. It then continued to be displayed until the next EMA was triggered (i.e., for 75 minutes), except if the participant actively dismissed it. With these adaptations in mind, both the schedule and the number of EMAs that were offered in a day varied between participants as well as between days.

3 DATA COLLECTION

3.1 Participants and Duration of the Study

Dutch and Slovenian speaking volunteers were recruited in research institutions in Belgium and Slovenia. The study was approved by the Commission of Medical Ethics of the Ghent University Hospital, Belgium (no. EC/2019/1091) and the Ethics Committee of the Faculty of Arts at the University of Ljubljana, Slovenia (no. 168-2019). The main part of the data collection ran from October 2020 to July 2021.

The main part of the study included 56 participants, of whom $N = 55$ filled in an initial online survey which asked about demographic and health behaviour information. Of these, 26 were women and 29 were men and 29 were residing in Belgium, while 26 lived in Slovenia. Their mean age was 34.9 years and ranged from 24 years to 63 years ($SD = 9.7$ years). They were all employed in research institutions, but held various positions, such as PhD students, employees in administration, and tenured professors.

3.2 Adherence

As the EMA schedule was variable due to customizations described in Section 2.1, there was no single approach to calculate participants' adherence. Only the EMA sessions that elicited a response from participants were recorded. This means that it was not possible to calculate adherence with a simple approach by comparing the number of offered to the number of filled out questionnaires. On the other hand, getting a response to an EMA prompt from a participant did not guarantee a completed EMA session (a full questionnaire), so we classified them first according to their completion.

Out of 6639 EMA sessions that participants initiated in total, around 14.8 % were merely short indications of either a finished working day or a non-working day. Another 4.2 % represented sessions that were interrupted: either the participants started answering but cut it short before completing the full EMA session, or the evening session was expired to make room for a morning EMA session. Most (81.0 %) of the EMA sessions that were initiated represented true questionnaires, which were also completed. These were considered in the following analysis as true EMA sessions.

In their participation period, participants finished more than 96 EMA sessions on average ($\mu = 96.3$), but this varied quite extensively from 50 to 152 ($SD = 21.0$). To consider the effect of gender, age, and country on the number of finished EMA sessions, a linear regression model was built. No differences in the number of completed EMAs were found between the two genders ($t = -1.18$ with male coded as 1, $p = 0.244$) and age was not a statistically significant predictor ($t = 1.34$, $p = 0.186$). Slovenian participants answered 11.9 more EMA sessions on average ($t = 2.18$ with Slovenian coded as 1, $p = 0.034$).

Next, the workday and the evening EMAs were considered separately. As described in Section 2.1, they consisted of different number of items (and also different content) and their schedules were qualitatively different.

3.2.1 Workday EMA. While the workday EMAs appeared every 90 minutes *on average*, their actual triggering time was dependent on an individual participant's preset habits and actual behaviour, such as postponing (dismissing) notifications and responding to them with varying latency. The bulk of the workday EMAs were filled in within approximately 90 minutes, but longer durations between subsequent workday EMA sessions were not uncommon, some several hours long.

The median time difference between two subsequent workday EMA sessions was 93 minutes ($\mu = 102$, $SD = 43$), with more than 80 % differences falling below two hours. Indeed, the median time difference of 90 minutes was typical for *most* of the participants, but there were also significant variations between them, as shown in Figure 1. Their individual median times were not related, however, to their gender ($t = -1.08$, $p = 0.285$), age ($t = -0.48$, $p = 0.632$), or language ($t = -0.36$, $p = 0.722$) in a linear regression model.

3.2.2 Evening EMA. Since the evening EMA was longer than the workday EMAs and it came outside of working hours, adherence was considered separately. Any day that a true EMA session was completed was checked for completion of the evening EMA and the ratio of these days was calculated.

On average, participants filled in the evening EMA on more than 90 % of their days of active participation (the median ratio was 0.93). They were more heterogeneous on this measure, however, since there were some that filled in the evening EMA less than 70 % of the days, as shown in Figure 2 As before, these differences could not be explained by their gender ($t = -0.98$, $p = 0.330$), age ($t = -0.43$, $p = 0.673$), or language ($t = -1.59$, $p = 0.118$) in a linear regression model.

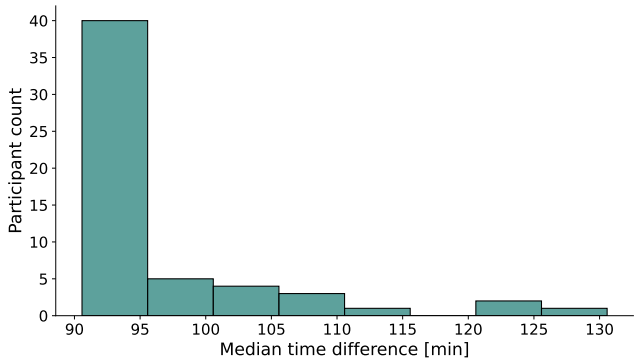

Figure 1: Distribution of median time differences between subsequent workday EMA sessions per participant.

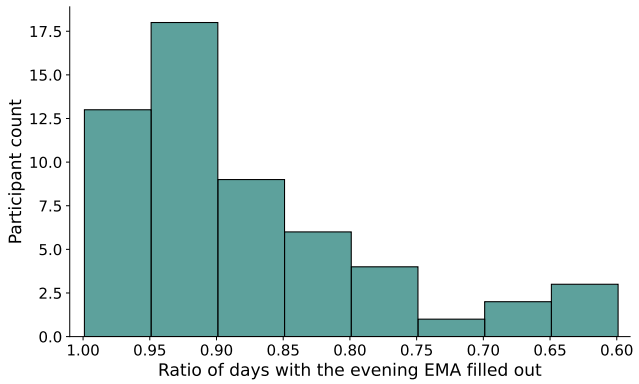

Figure 2: Distribution of the ratio of the days in which the evening EMA was filled out (calculated per participant).

3.3 Participants' Feedback

In a debriefing session, participants were asked for positive and negative feedback regarding their participation in the study.

Most negative remarks about EMAs were about participants needing to get used to them and that they were sometimes difficult to answer on a 90-minute basis. For example, it was not completely clear how to describe coping with stressful situations when none occurred on that workday. Some questions confused participants with similar terms, such as *stress* and *tension* in the Dutch translation of the Stress Appraisal Measures [10]. They also expressed frustration with no option to go back to already answered items, which was a purposeful design decision to encourage quick and instinctive answers.

On the positive side, participants pointed out that EMAs encouraged self-reflection about work and their responses to stressful situations. This sometimes helped them frame negative experiences as challenges or furthered their awareness of many different tasks in the workday and their various activities. They also liked the feedback that was displayed every morning after the first EMA session. It informed them of how many EMAs they had responded to in the previous day and encouraged continued participation.

In general, participants reported that the EMAs did not present an excessive burden. They tried to respond as soon as possible, but most often they postponed their response because of meetings or calls and focused or laboratory work. They were also less responsive during days that demanded harder work, such as before deadlines.

The biggest concern with the Android application was battery drain which was increased because of multiple sensors running continuously. This was alleviated by warning participants about it in the briefing session, but some wanted some sensors disabled (e.g., accelerometer). The battery optimization functions sometimes prevented the application to trigger a notification, but this was usually resolved by unlocking the phone and looking for the EMA actively.

4 COMPARISON WITH TURNOUT BURNOUT PROJECT

Turnout Burnout was a project that ran in 2013 and was funded by the European Institute of Innovation & Technology, Information and Communication Technologies Labs (EIT ICT Labs), now EIT Digital. The project included six European partners where one of the objectives was to detect early sings of occupational stress and consequent burnout. Since both the research topic as well as methodology were similar to the STRAW project, the results and researchers' experiences can be compared.

4.1 Study Design

In comparison to the STRAW, this study was longer and ran for 8 weeks during November and December 2013. Participants signed an informed consent and were recruited based on a presentation of the project background, its objectives and the study design, that was held at each organisation. Approximately twice as many participants took part in the presentation in comparison to the final participants in the study. The participants that refused to take part in the study cited the need to change their smartphone as the primary impediment to participation. The study protocol defined that in order to take part, each participant had to ensure that the project-provided phone became the primary smartphone, with the participants' own subscriber identity module (SIM) inserted into it. The users were instructed to handle their project-provided phone as they would their own phone to ensure the collected data reflects participants' usual behaviour.

The phone (Samsung Galaxy S3 mini, 32GB) was delivered preconfigured to the participants, with a purpose developed sensing app. Ready-made frameworks (such as AWARE, used in the STRAW project) were not widely available at the time. The app started automatically on working days only (Monday through Friday), without interaction from the user, and ran in the background, continuously collecting sensor data.

In addition to the (objective) sensor data (described in [6], [9] and [4]), the app also collected subjective variables. In order to understand users' mood and stress levels, the app prompted users to fill in a questionnaire at three different times of the day: at 9:00 (at the beginning of the work hours), at 14:00 (after lunch break) and at 17:00 (at the end of the work hours). As with the STRAW approach, the questionnaires appeared automatically, and the user had the option to answer the questions or snooze the questionnaire for later.

The questionnaire was derived from two validated questionnaires of burnout and mood.

4.2 Study Participants and Adherence Levels

Turnout Burnout participants were somewhat older (μ = 37.46 years, SD = 7.15 years) in comparison to the STRAW participants, with lower number of female participants (40 % vs 46 % of women). One third of participants had a graduate degree (33.3 %), while 36.7 % had an undergraduate degree, with the rest having a high-school diploma [8].

Analysis of adherence revealed that 1455 questionnaires were completed by the participants, representing a response rate of 79.97 %. This cannot be directly compared to the STRAW study, since the schedule in the latter was not fixed and therefore the number of questionnaires offered depended on participants' behaviour.

4.3 Participants' Feedback

While no formal feedback was collected at the end of the study, informal feedback provided to the organisers of the study touched upon several aspects.

The main negative remark was related the battery life of the device. Even though the app was optimised in this respect, the trade off between battery life and frequency of data collection did not meet participants' expectations, where some of them reported having to charge the phone before the end of the day. While the majority of participants did not perceive the questionnaires as a significant burden, a small number of participants suggested using a lower number of questions.

On the positive note, the majority of the participants found the study interesting and thought a service that would provide insights into their work life would have been useful. It is interesting to note that some of the participants were also driven by altruistic motivation by dedicating their time and effort in contributing to a scientific project.

5 CONCLUSIONS AND LESSONS LEARNED

The studies described in this paper had a longer data collection period or a larger number of questionnaire sessions than comparable diary studies [cf. 1]. While the adherence levels are difficult to compare due to heterogeneous methodologies, in a similar but shorter study of physicians and nurses [11], 7.8 % and 17.5 % of responses, respectively, were incomplete, compared to the 4.2 % in the STRAW study. On the other hand, 75.8 % and 69.3 % of responses, respectively, were complete, compared to the 80 % in the Turnout Burnout study.

Experiences of data collection in both studies thus indicate that long EMA studies even with a large amount of questionnaires can be successfully carried out. After the initial commitment, none of the participants dropped out. Active involvement of researchers in the data collection process proved to be crucial in motivating participants to start and keep up their participation in the study.

Establishing good rapport in the briefing session was very important, both to give clear instructions for following the somewhat complex protocol and to make sure the participants understand what types of data are collected and how their privacy is protected. Since a lot of data was gathered in the process and it came from different sources, it was essential that participants understand what they were revealing on top of their declared answers to EMAs.

Equally valuable was continuous contact between researchers and participants. The former made clear they were always available for any questions and tried to solve issues as soon as they appeared. They also actively sought confirmation from participants that everything is going well and gave feedback on how things are running server-side where data collection could be monitored. If participants so desired, they received daily reminders to transfer the wristband data by the researchers.

Providing participants with a dedicated device or using their own should be decided based on the research question of interest. But since participants heavily relied on their personal devices, it was important in both cases to set expectations regarding battery life and offer alternative solutions to decrease the chance of drop out.

ACKNOWLEDGMENTS

The authors acknowledge the project STRAW was financially supported by the Slovenian Research Agency (project reference N2-0081) and the Research Foundation–Flanders of Belgium (project no. G.0318.18N).

REFERENCES

[1] Larissa Beattie and Barbara Griffin. 2014. Day-level fluctuations in stress and engagement in response to workplace incivility: A diary study. *Work Stress* 28, 2 (April 2014), 1–19. https://doi.org/10.1080/02678373.2014.898712

[2] Larissa Bolliger, Junoš Lukan, Mitja Luštrek, Dirk De Bacquer, and Els Clays. 2020. Protocol of the Stress at Work (STRAW) Project: How to Disentangle Day-to-Day Occupational Stress among Academics Based on EMA, Physiological Data, and Smartphone Sensor and Usage Data. *International Journal of Environmental Research and Public Health* 17, 23 (nov 2020), 8835. https://doi.org/10.3390/ijerph17238835

[3] Eurofound. 2017. *Sixth European Working Conditions Survey – Overview report (2017 update).* Publications Office of the European Union, Luxembourg. https://doi.org/10.2806/422172

[4] Raihana Ferdous, Venet Osmani, Jessica Beltrán Márquez, and Oscar Mayora. 2015. Investigating correlation between verbal interactions and perceived stress. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* IEEE, Milan, Italy, 1612–1615. https://doi.org/10.1109/embc.2015.7318683

[5] Denzil Ferreira, Vassilis Kostakos, and Anind K. Dey. 2015. AWARE: Mobile Context Instrumentation Framework. *Frontiers in ICT* 2, 6 (2015), 1–9. https://doi.org/10.3389/fict.2015.00006

[6] Enrique Garcia-Ceja, Venet Osmani, and Oscar Mayora. 2016. Automatic Stress Detection in Working Environments From Smartphones' Accelerometer Data: A First Step. *IEEE Journal of Biomedical and Health Informatics* 20, 4 (jul 2016), 1053–1060. https://doi.org/10.1109/jbhi.2015.2446195

[7] Gillian H. Ice and Gary D. James. 2006. Conducting a field study of stress. In *Measuring Stress in Humans*, Gillian H. Ice and Gary D. James (Eds.). Cambridge University Press, Cambridge, UK, Chapter 1, 3–24.

[8] Alban Maxhuni, Pablo Hernandez-Leal, Eduardo F. Morales, L. Enrique Sucar, Venet Osmani, and Oscar Mayora. 2021. Unobtrusive Stress Assessment Using Smartphones. *IEEE Transactions on Mobile Computing* 20, 6 (jun 2021), 2313–2325. https://doi.org/10.1109/tmc.2020.2974834

[9] Venet Osmani, Raihana Ferdous, and Oscar Mayora. 2015. Smartphone app usage as a predictor of perceived stress levels at workplace. In *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare.* IEEE, ICST, Istanbul, Turkey, 225–228. https://doi.org/10.4108/icst.pervasivehealth.2015.260192

[10] Edward J. Peacock and Paul T. P. Wong. 1990. The stress appraisal measure (SAM): A multidimensional approach to cognitive appraisal. *Stress Medicine* 6, 3 (jul 1990), 227–236. https://doi.org/10.1002/smi.2460060308

[11] Thomas Rutledge, Erin Stucky, Adrian Dollarhide, Martha Shively, Sonia Jain, Tanya Wolfson, Matthew B. Weinger, and Timothy Dresselhaus. 2009. A real-time assessment of work stress in physicians and nurses. *Health Psychology* 28, 2 (March 2009), 194–200. https://doi.org/10.1037/a0013145