Smartphone Based Stress Prediction

Thomas Stütz[1], Thomas Kowar[1], Michael Kager[1], Martin Tiefengrabner[1], Markus Stuppner[2], Jens Blechert[2], Frank H. Wilhelm[2], and Simon Ginzinger[1](✉)

[1] Department of Multimedia Technology, University of Applied Sciences Salzburg, Campus Urstein Süd 1, 5412 Puch/Salzburg, Austria
{thomas.stuetz,thomas.kowar,michael.kager, martin.tiefengrabner,simon.ginzinger}@fh-salzburg.ac.at
http://fh-salzburg.ac.at/mmt

[2] Department of Psychology, University of Salzburg, Hellbrunnerstrasse 34, 5020 Salzburg, Austria
{markus.stuppner,jens.blechert,frank.wilhelm}@sbg.ac.at
http://uni-salzburg.ac.at/psychologie

Abstract. Smartphone usage has tremendously increased and most users keep their smartphones close throughout the day. Smartphones have a broad variety of sensors, that could automatically map and track the user's life and behaviour. In this work we investigate whether automatically collected smartphone usage and sensor data can be employed to predict the experienced stress levels of a user using a customized brief version of the Perceived Stress Scale (PSS). To that end we have conducted a user study in which smartphone data and stress (as measured by the PSS seven times a day) were recorded for two weeks. We found significant correlations between stress scores and smartphone usage as well as sensor data, pointing to innovative ways for automatic stress measurements via smartphone technology. Stress is a prevalent risk factor for multiple diseases. Thus accurate and efficient prediction of stress levels could provide means for targeted prevention and intervention.

Keywords: Stress · Prediction · Smartphone sensing · Data analysis · Field study · Observational study

1 Introduction

Stress has become a major concern, as stress-related diseases cause a decrease in life expectancy and in the quality of life, as well as an increase in the number of sick leaves. Stress is a risk factor for a multitude of diseases, also because stress is closely linked to unhealthy behaviour, such as excessive consumption of food and alcoholic beverages [1,21]. Accurate and efficient assessment of unhealthy stress exposure could greatly help to administer timely and appropriate support measures. Smartphones can offer a deep insight in the user's daily life, due to their multitude of sensors, their frequent use in everyday life and their constant

F. Ricci et al. (Eds.): UMAP 2015, LNCS 9146, pp. 240–251, 2015.
DOI: 10.1007/978-3-319-20267-9_20

proximity to the user. Therefore it would be desirable to develop stress indices based on smartphone usage and sensor data.

If subjective stress levels were predictable by these data, health-related smartphone apps can automatically take appropriate actions, such as a notification of the responsible physician or prompts to the user to reduce stress through other means. The main objective of the study is to research the connection between perceived stress and smartphone usage and sensor data. More specifically the objective is to answer the question whether there is a significant correlation between perceived stress scores and smartphone usage and sensor data.

The presented study was part of a broader study on the relation between stress and diet. The study was formally approved by the ethics committee of the University of Salzburg.

A brief overview of related work is given in section 2. The presentation of the study is structured as recommended in [22]. The context of the study is explained in section 3. In section 4 the methodology is presented and in section 5 the results are given, which are discussed in section 6. Finally, we conclude in section 7.

2 Related Work

Initial results on stress detection with smartphones are presented in [4]. Mobile phone usage data of a stressful (exam period) and a normal week is compared in seven subjects. In [20] wrist sensor data and smart phone data has been used to predict high and low levels of stress, as measured by the full PSS questionnaire [7]. Sensor and smart phone usage data of 18 persons was collected for 5 days and 75% accuracy was reported for binary classification (high / low stress). Stress prediction on the basis of sophisticated audio processing is presented in [17]. Stress levels were measured with a GSR (galvanic skin response sensor). In [23] smartphone usage and sensor data collected by a continuous sensing app has been shown to correlate with academic performance and various psychological scores, namely PHQ-9 for depression, PSS for stress, flourishing scale, and UCLA loneliness scale. The study with 48 participants has a broad spectrum and does not focus on stress alone. Compared to our study, stress is measured only before and after the study with the full PSS questionnaire [7].

The influence of weather and mobile phone usage on perceived stress is investigated in [5]. Phone usage data consisted of call logs (no audio properties), sms logs (no content based features), and proximity data (other Bluetooth devices). Stress levels were measured by one direct question with a 7-point Likert scale to answer. Data was collected from 117 persons for almost half a year. For binary classification (stress / no stress) an 72.28% accuracy is reported.

A mobile phone app for depression is presented and analysed in [6]. The application provided self-monitoring via questionnaires and tailored, real-time feedback and intervention. Somewhat similar work for bipolar disorder is presented in [3,11]. The application of smartphones (and processed sensor data) for large scale behaviour change interventions is discussed in [15].

3 Study Context

The study was performed at the department of Psychology at the University of Salzburg and the participants were students of psychology. The smartphone app (TheStressCollector) was developed at the department of MultiMediaTechnolgy at the Salzburg University of Applied Sciences. The study was joint work of the two departments. In the study two smartphone applications were employed. The smartphone app TheStressCollector (TSC) was responsible for logging usage and sensor data, while the commercial framework movisensSX [19] was used to administer the questionnaires. MovisensXS allows non-programming experts to conduct questionnaire-based studies for Android smartphones.

3.1 Mobile App for Smartphone Usage and Sensor Logging

The smartphone app TheStressCollector (TSC), once installed on an Android smartphone, runs permanently in the background and periodically collects smartphone usage and sensor data and uploads the data to a database server (PHP, MySQL). For participants who want to avoid mobile data usage an option to only synchronize via WiFi network is offered.

All collected data were fully anonymized, i.e. the subjects were identified by a random ID. Furthermore only data items that did not allow to trace back a person's identity have been collected.

In the following we give a brief overview of the collected data items (see Fig. 1 as well):

Activity: TSC uses the activity recognition provided by Google Play Services [2]. In the following we list the different activities detected, which were classified into physical versus non-physical activities.

- ON_FOOT: The device is on a user who is walking or running. This is classified as *physical.*
- ON_BICYCLE: The device is on a bicycle. This is classified as *physical.*

Fig. 1. Smartphone Usage and Sensor Logging

- IN_VEHICLE: The device is in a vehicle. This is classified as *non-physical*.
- TILTING: The device angle relative to gravity changed significantly. This is most often the case when the user is using the phone at the moment. This is classified as *non-physical*
- STILL: The device is not moving. This is classified as *non-physical*.
- UNKNOWN: It is not possible to detect the current activity. This is classified as *physical*.

The activity detection interval is set to 20 minutes. More frequent measurements may occur in case another application request more frequent activity updates. Less frequent measurements may occur in case the device is in the STILL state.

App Usage: Due to privacy concerns the TSC app does not save the exact package names of executed apps, but categorizes them into one of the six rough categories Information, System, Health, Social, Entertainment or Work.

Table 1. App categories of TSC and the Google Play Store

Category	**Google Play Store Categories**
Information	Transportation, Weather, Travel & Local, News & Magazines, Shopping and Lifestyle
System	Libraries & Demo, Personalization, Live Wallpapers, Tools and Widgets
Health	Sports, Medical and Health & Fitness
Entertainment	Photography, Entertainment, Music & Audio, Media & Video, Books & Reference and 19 game categories
Social	Communication and Social
Work	Business, Productivity, Finance and Education

Network Traffic: As a measurement of general smartphone usage intensity, the amount of network traffic (in terms of bytes) was logged as well. TSC splits the traffic into the categories 'Received' and 'Transmitted' as well as 'Mobile Network' and 'Wireless Network'.

Reboot Activity: TSC logs power on and power off events.

Calls: If a call gets active, the microphone input is processed in real-time. The power levels in *dB* are measured during the call. Measurements were taken every 300 milliseconds. After the call has finished, the minimum, maximum and mean *dB*-values are saved. Furthermore the call duration is logged. Calls within working hours are counted separately as well (call working count).

Light: The environment brightness as measured by the light sensor was measured. Measurements were taken whenever the screen was on, i.e. when the user was interacting with the smartphone (to avoid meaningless measurements when the smartphone is stored e.g. in a pocket). Each measurement has been classified into 'light' or 'dark' and the frequency summarized in 'light session count' and dark 'session count'.

Messages: TSC logs time stamps for every SMS and MMS received on the smartphone, but neither the content nor the receiver (no privacy incriminating information). Messengers like WhatsApp, Google Hangouts or Facebook Messenger are covered by the app usage tracking and categorized as 'Social'.

Noise exposure: Noise exposure was logged periodically, i.e., the minimum, maximum and mean power in dB of a 20 seconds sample (with a sampling rate of 100 milliseconds) was computed every 20 minutes.

Screen Activity: User session duration was logged. A user session starts when the device screen is powered on and ends if the screen goes off. User sessions during working time are counted separately as well (user session working time) User session percent summarizes the ratio of active smartphone usage and no usage.

4 Methods

We conducted a longitudinal, ambulatory, observational study on smartphone usage, stress, and eating behaviour. The study was conducted in 2014 in the summer semester and included one week in mid-semester (assumed not very stressful) and one week at the end of the semester (assumed stressful due to end-term examinations).

There were no exclusion criteria for participation in the study. Participants installed the app on their private smartphones. Seven times a day questionnaires including the PSS had to be answered, i.e. at 8:00, 9:00, 12:00, 15:00, 18:00, 21:00, and 22:00. Only in the morning a questionnaire on sleeping behaviour was scheduled. At 22:00 an extensive set of questions on eating and stress was scheduled. Throughout the day a short questionnaire on food and drink consumption was asked repeatedly. During the study, the test users had access to a website which provided information about the study and clear instructions regarding what to do before and after the two study weeks. For purpose of installation on the participants' smartphones the TSC app was made available in the Google Play Store.

4.1 Outcome Measures for Perceived Stress

The perceived stress score (PSS) has been introduced in [7]. The original PSS questionnaire consists of 14 questions. For the study the original questionnaire was considered too time-consuming and thus an abbreviated four question version of the PSS was employed. Internal consistency (based on averaged items over all assessment points of week 1) was good (.831). The following four questions were asked repeatedly:

1. Since the last entry: Has something unexpected upset you?
2. Do you feel nervous and stressed?

3. Since the last entry: Have you felt that you were on top of things?
4. Do you feel that you can successfully cope with all upcoming tasks and problems at the moment?

Questions 1, 2, and 4 could be answered on a 7-level Likert scale, while question 3 could be answered on 8-level Likert scale (which was a human error in the setup of the study). Thus for each question the answer is a number in the range of 1 (not at all) to 7 or 8 (extremely). The PSS is simply computed by adding the scores of answers 1 and 2 and the inverted scores of answers 3 and 4. Thus high PSS values (max = 29) correspond to high level of perceived stress and low PSS values (min = 4) correspond to a low level of perceived stress.

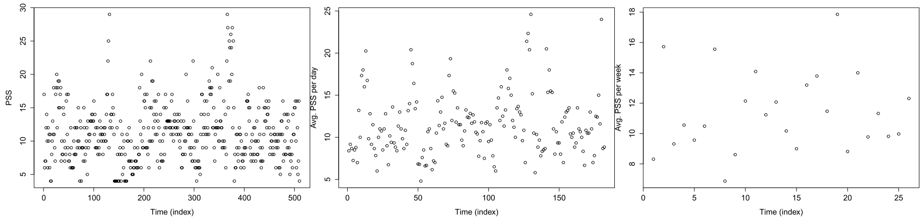
4.2 Methods for Data Analysis

Data analysis was conducted in R (version 3.1.2 for Windows). The Pearson correlation coefficient and its significance level [16] were computed with the R-function `cor.test`.

Preliminary results on prediction were generated with the data mining software WEKA (version 3.7 for Windows) [13,24]. The WEKA software was used to evaluate the prediction performance of standard machine learning tools on the collected data. Specifically linear regression and random forests [24] were employed with 5-fold and 10-fold cross validation.

5 Results

PSS and smartphone usage and sensor data from 15 participants (12 women and 3 men) were collected. PSS measurements resulted in 104898 data points; 967 PSS measurements were obtained in the first week, and 754 PSS measurements in the second week. In the first week the average response rate results to 85% with a standard devision of 12%. In the second week the average response rate results to 88% with a standard devision of 13%. The average PSS value in the first week (10.3) wasn slighly lower as compared to the second week (11.0). Detailed information on the participants is summarized in table 2. Initially 28 participants had been recruited for the study. The decrease to 15 participants with associated data ist due to the entry of mismatching identification codes in the TSC and the MovisensSX app, personal smartphones with insufficient free memory (therefore only the MovisensSX app was installed), and privacy concerns. Out of the 15 participants of the first week four participants did not return to the second week. All PSS data from both weeks used in further data analysis is shown in Figure 2. The PSS data is ordered by the time it was received at the server. Table 3 summarizes the correlation and associated p-values of single PSS scores with different collected data items. For the correlations and p-values in table 4 the seven single PSS scores of each day were averaged, while for table 5 all PSS scores of each week were averaged. Figures 3 and 4 show plots of the most significant data items versus average PSS scores.



(a) (b) (c)

Fig. 2. PSS values (a), average PSS per day (b), and average PSS per week (c)

Table 2. Participant details

Id	Gender	BMI	Age(years)	Height(cm)	Weight(kg)	Smoker	Diet	Edu
3	female	22.0	23	165	60	no	non-veg.	univ.
5	female	19.1	23	168	54	no	non-veg.	univ.
8	female	21.3	24	162	56	yes	non-veg.	univ.
9	female	25.0	24	165	68	no	non-veg.	univ.
10	female	19.4	22	173	58	no	veg.	univ.
11	female	19.1	21	155	46	no	non-veg.	A-level
12	female	22.9	23	162	60	no	non-veg.	univ.
13	female	19.6	27	178	62	no	non-veg.	univ.
17	female	21.8	24	163	58	no	non-veg.	univ.
18	male	23.3	26	192	86	no	non-veg.	univ.
19	female	24.9	21	169	71	no	non-veg.	A-level
23	female	22.8	27	169	65	no	non-veg.	univ.
25	female	20.2	24	165	55	no	veg.	univ.
26	male	24.2	24	183	81	no	non-veg.	A-level
27	female	21.5	33	178	68	no	non-veg.	univ.

Table 3. Single perceived stress score and Pearson correlation with the 10 most significant data items

Automatically logged data item	Pearson cor	Pearson p-value
noisiness min	0.2024	0.00002
social count	0.1135	0.01033
app mean time cat health	-0.1119	0.01350
call count	0.1022	0.02195
app mean time cat system	-0.1018	0.02465
wifi received	0.0930	0.03578
app sum time cat health	-0.0947	0.03673
noisiness mean	0.0963	0.04309
lighter sessions count	0.1275	0.05400
user session count	0.0958	0.10099

The aim is to predict excessive levels of stress, thus we have employed the collected data as data set for machine learning algorithms. The results, i.e., the correlation of the predicted stress score and the actual stress score, are

Table 4. Daily average of perceived stress scores and Pearson correlation with the 10 most significant data items

Automatically logged data item	Pearson cor	Pearson p-value
noisiness mean	0.3214	0.00001
power on count	0.2841	0.00010
noisiness min	0.2579	0.00047
call max db	0.2840	0.00111
call mean db	0.2833	0.00114
light changes count	0.3105	0.00479
app mean time cat work	0.2082	0.00568
call count	0.2022	0.00730
call sum time	0.2256	0.01014
lighter sessions count	0.2706	0.01456

Table 5. Weekly average of the perceived stress scores and Pearson correlation with the 10 most significant data items

Automatically logged data item	Pearson cor	Pearson p-value
user session percent	0.6506	0.00862
user session sum time	0.6503	0.00867
activity difference (physical - non-physical)	0.6758	0.01124
user session sum working time	0.6065	0.01652
activity nonphysical count	-0.6388	0.01875
user session max time	0.5638	0.02859
power on count	0.3527	0.07718
user session mean time	0.4521	0.09063
noisiness mean	0.3230	0.10755
call working count	0.3113	0.12988

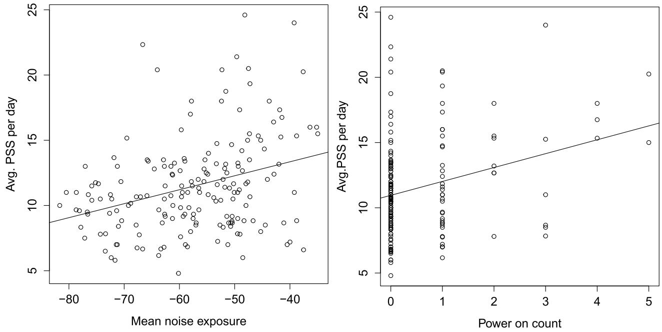


Fig. 3. Daily PSS averages and selected smartphone data

summarized in tables 6 and 7 for linear regression and random forests. The quality of the predicted stress levels (predicted by WEKA's machine learning algorithms) was evaluated in terms mean absolute error (MAE) to and Pearson

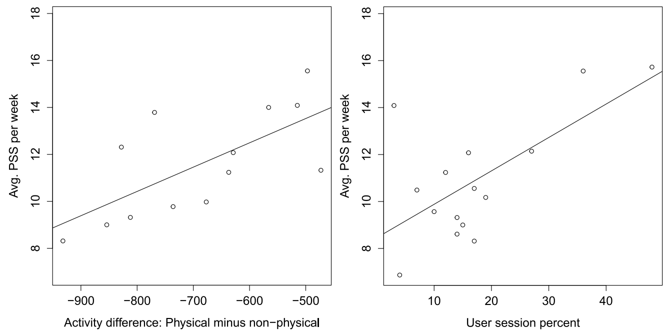


Fig. 4. Weekly PSS averages and selected smartphone data

Table 6. Results for predicting the PSS with linear regression

Dataset	Fold 5			Fold 10		
	MAE	cor	p-value	MAE	cor	p-value
PSS, single (510 instances)	3.7	0.08	0.07448	3.7	0.07	0.10840
PSS, daily average (182 instances)	3.5	0.14	0.06006	3.5	0.14	0.08606
Week (26 instances)	6.7	-0.15	0.45770	9.2	-0.19	0.34750

Table 7. Results for predicting the PSS with random forests

Dataset	Fold 5			Fold 10		
	MAE	cor	p-value	MAE	cor	p-value
PSS, single (510 instances)	3.3	0.15	0.00074	3.3	0.16	0.00043
PSS, daily average (182 instances)	2.4	0.4	1.02e-08	2.4	0.45	2.28e-10
Week (26 instances)	2.0	0.03	0.87530	1.9	0.16	0.42130

correlation with the actual stress levels (using cross validation). We also give the p-value of the Pearson correlation. The correlation plot for the most significant predictor is shown in Figure 5

6 Discussion

The results show significant correlations between PSS and smartphone usage and sensor data. Interestingly the highest correlation coefficient value is found with the weekly PSS average, which indicates that longer periods of high stress exposure could be identified better than short periods. However, the most significant result is not found for weekly average of the PSS, because too few data points are available. The most significant result is found for the daily averages of PSS and the average noise exposure (noisiness mean).

Overall we find significant correlations (p-value < 0.05) for single PSS, daily averaged PSS, and weekly averaged PSS.

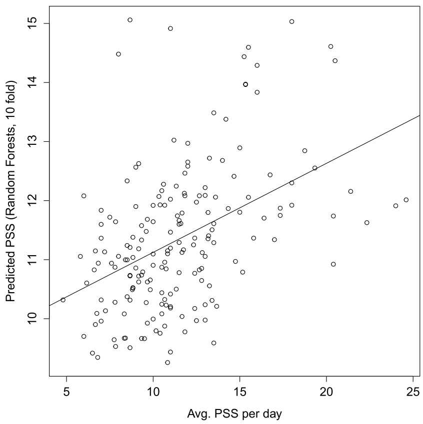


Fig. 5. Correlation Plot for most significant predictor

For a single PSS measurement, recent noise exposure (noisiness min) and social contacts (social count) are most significant influences. Noise exposure is also the most significant influence for the daily average of PSS values. The negative influence of noise and health has been previously reported [12,18]. Social interactions and stress are discussed in [8] and proprieties of conversations (duration, frequency) have been found to correlate negatively with PSS [23].

The number of times a user has pressed the power on button of his smartphone also shows high correlation with the daily average of PSS (power on count).

For the weekly average of PSS correlation with active usage of the smartphone is most significant. Interestingly there is also a significant correlation with physical activity. This seems to contradict previous findings on activity and mental well-being. Physical activity and mental well-being have been researched in [9,10,14], assessing a positive influence of physical activity on mental well-being. However, we believe that our findings reflect the circumstance that students have to go the exams more often in the second week (walking has been counted as a physical activity).

The prediction results show potential, but are not yet very accurate. We believe a bigger data set will lead to better results and that predictions targeted for individual persons could perform better. Our results foster the hope that given sufficient training data automatic stress predictors are feasible.

Despite the small number of participants we achieve highly significant correlations, even as compared to larger studies [23]. Most probably this is due to the focus on stress and the associated evaluation of PSS several times a day.

In future work we plan to extend and improve our app, especially with respect to localization analysis, conversation analysis and sleep. We hope that further improvements in sensor data processing bring us closer to the goal of fully automatic stress prediction.

7 Conclusion

We have conducted an observational study on stress and smartphone usage and sensor data. Our results show significant correlations between stress and smartphone data and outperform previously reported significance levels. Our study encourages future work and further studies on smartphone based stress prediction.

Acknowledgments. This work was supported by the Austrian research promotion agency (FFG), project number 839076.

References

1. Adam, T.C., Epel, E.S.: Stress, eating and the reward system. Physiology & behavior **91**(4), 449–458 (2007)
2. Android: Google play services (August 2014). https://developer.android.com/google/play-services/index.html
3. Bardram, J.E., Frost, M., Szántó, K., Marcu, G.: The monarca self-assessment system: a persuasive personal monitoring system for bipolar patients. In: Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, pp. 21–30. ACM (2012)
4. Bauer, G., Lukowicz, P.: Can smartphones detect stress-related changes in the behaviour of individuals? In: IEEE International Conference on Pervasive Computing and Communications Workshops, pp. 423–426. IEEE (2012)
5. Bogomolov, A., Lepri, B., Ferron, M., Pianesi, F., Pentland, A.S.: Daily stress recognition from mobile phone data, weather conditions and individual traits. In: Proceedings of the ACM International Conference on Multimedia, pp. 477–486. ACM (2014)
6. Burns, M.N., Begale, M., Duffecy, J., Gergle, D., Karr, C.J., Giangrande, E., Mohr, D.C.: Harnessing context sensing to develop a mobile intervention for depression. Journal of medical Internet research **13**(3) (2011)
7. Cohen, S., Kamarck, T., Mermelstein, R.: A global measure of perceived stress. Journal of health and social behavior, 385–396 (1983)
8. Coyne, J.C., Downey, G.: Social factors and psychopathology: Stress, social support, and coping processes. Annual review of psychology **42**(1), 401–425 (1991)
9. Dunn, A.L., Trivedi, M.H., O'Neal, H.A.: Physical activity dose-response effects on outcomes of depression and anxiety. Medicine & Science in Sports & Exercise **33**, S587–97 (2001)
10. Fox, K.R.: The influence of physical activity on mental well-being. Public health nutrition **2**(3a), 411–418 (1999)
11. Frost, M., Doryab, A., Faurholt-Jepsen, M., Kessing, L.V., Bardram, J.E.: Supporting disease insight through data analysis: refinements of the monarca self-assessment system. In: ACM international joint conference on Pervasive and ubiquitous computing, pp. 133–142. ACM (2013)
12. Haines, M.M., Stansfeld, S.A., Job, R.S., Berglund, B., Head, J.: Chronic aircraft noise exposure, stress responses, mental health and cognitive performance in school children. Psychological medicine **31**(02), 265–277 (2001)

13. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: An update. SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
14. Hansmann, R., Hug, S.M., Seeland, K.: Restoration and stress relief through physical activities in forests and parks. Urban Forestry & Urban Greening **6**(4), 213–225 (2007)
15. Lathia, N., Pejovic, V., Rachuri, K.K., Mascolo, C., Musolesi, M., Rentfrow, P.J.: Smartphones for large-scale behavior change interventions. IEEE Pervasive Computing **3**, 66–73 (2013)
16. Lee Rodgers, J., Nicewander, W.A.: Thirteen ways to look at the correlation coefficient. The American Statistician **42**(1), 59–66 (1988)
17. Lu, H., Frauendorfer, D., Rabbi, M., Mast, M.S., Chittaranjan, G.T., Campbell, A.T., Gatica-Perez, D., Choudhury, T.: Stresssense: detecting stress in unconstrained acoustic environments using smartphones. In: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp. 351–360. ACM (2012)
18. Matheson, M., Stansfeld, S., Haines, M., et al.: The effects of chronic aircraft noise exposure on children's cognition and health: 3 field studies. Noise and Health **5**(19), 31 (2003)
19. movisensXS: experience sampling for android (August 2014). https://xs.movisens.com
20. Sano, A., Picard, R.W.: Stress recognition using wearable sensors and mobile phones. In: Humaine Association Conference on Affective Computing and Intelligent Interaction, pp. 671–676. IEEE (2013)
21. Schwabe, L., Dickinson, A., Wolf, O.T.: Stress, habits, and drug addiction: a psychoneuroendocrinological perspective. Experimental and clinical psychopharmacology **19**(1), 53 (2011)
22. Talmon, J., Ammenwerth, E., Brender, J., de Keizer, N., Nykänen, P., Rigby, M.: STARE-HI: Statement on reporting of evaluation studies in health informatics. International Journal of Medical Informatics **78**(1), 1–9 (2009)
23. Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., Zhou, X., Ben-Zeev, D., Campbell, A.T.: Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 3–14. ACM (2014)
24. Witten, I.H., Frank, E.: Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann (2005)