Design Decisions for a Real Time, Alcohol Craving Study Using Physio- and Psychological Measures

Hendrika G. van Lier[1(✉)], Mira Oberhagemann[1], Jessica D. Stroes[1], Niklas M. Enewoldsen[1], Marcel E. Pieterse[1], Jan Maarten C. Schraagen[1,2], Marloes G. Postel[1,3], Miriam M.R. Vollenbroek-Hutten[1,4], Hein A. de Haan[3], and Matthijs L. Noordzij[1]

[1] University of Twente, Postbus 217, 7500 AE Enschede, The Netherlands
h.g.vanlier@utwente.nl
[2] Netherlands Organization for Applied Scientific Research (TNO), Soesterberg, The Netherlands
[3] Tactus Addiction Treatment, Deventer, The Netherlands
[4] Ziekenhuis Groep Twente, ZGT Academy, Almelo, The Netherlands

Abstract. The current study was a pilot for an alcohol craving monitoring study with a biosensor (E4 wristband) and ecological momentary assessment (EMA) smartphone app. The E4 wristband was evaluated on compliance rates, usability, comfort and stigmatization. Two EMA methodologies (signal- and interval-contingent design) were compared on data variability, compliance and perceived burden. Results show that both EMA methodologies captured variability of craving and compliance rates were between medium to low. The perceived burden of the designs was high, in particular for the signal-contingent design. Participants wore the wristband ranging from occasionally to often and the usability was rated good. Many participants reported frequent questioning about the bracelet, which they indicated as positive. However, addicted individuals are expected not to appreciate this attention, we therefore propose to provide them with coping strategies. Efforts should be made to increase compliance, we therefore propose the interval contingent design with micro incentives.

Keywords: Biosensor · Ecological momentary assessment · Data variability · Compliance rates · Perceived burden · Usability · Wearing comfort · Stigmatization

1 Introduction

Alcohol abuse is currently the fourth leading unhealthy lifestyle behavior contributing to morbidity and mortality. Oinas-Kukkonen and Harjumaa [1] argue that in the case of (alcohol) addiction, persuasive systems should aim at reinforcing proper attitudes and making them easier to stick with even in challenging,

P.W. de Vries et al. (Eds.): PERSUASIVE 2017, LNCS 10171, pp. 3–15, 2017.
DOI: 10.1007/978-3-319-55134-0_1

spontaneous situations. In a similar vein, Rohsenow and Monti [2] state that prompting patients to mobilize their coping resources by making them aware of their craving in challenging situations may protect them from relapse.

Innovations in wearable technology offer new possibilities to determine these challenging craving situations, by continuous measurement of fluctuations in physiological parameters, not only in the lab setting, but also in the real world (e.g. [3]). Craving responses have been proven multiple times in lab settings. According to a meta-analysis [4], alcoholics evidently show heightened physiological responses (heart rate, electrodermal activity) and psychological response (self-reported craving) to alcohol cues (e.g., pictures of beer). Yet, whether these effects are transferable from the lab setting to continuously measuring in the real world, the "wild", has not been investigated, therefore this pilot and eventually a monitoring study is performed. The lack of prior studies impedes making informed study design decisions. This study pilots the design decisions of an alcohol craving monitoring study using a smartphone app and a wearable biosensor. Using an app to administer questions is known to increase the compliance and to reduce the perceived burden [5]. The questionnaires were administered according to a Ecological Momentary Assessment design (EMA, [3]). EMA studies are repeated measurements of participants experiences and behavior in real time and in their natural environment [6].

A possible limitation of measuring in the "wild" is the completeness of both the EMA and wearable and data. Since the completeness of the data relies on the compliance of respondents and the functioning of the technology, missing values are likely to occur [6]. Additionally, if the wearable biosensor is not perceived as usable and comfortable, this could be a reason to stop wearing the sensor over time leading to an increase in the number of missing values. Furthermore, because the target population are alcoholics who are known to value anonymity [7], wearing a biosensor for an alcohol craving study might establish a feeling of stigmatization. A possible drawback of using a repeated intensive assessment like EMA is the high burden which might discourage participation. This can result in a sampling bias where only participants with certain personality factors and high motivation enter or complete the assessment [8]. This might also impose an compliance problem for the sample in general, since addicts are not known for their motivation, responsibility and compliance to schedules and instructions [6].

Two design decisions, the EMA cue type and the user acceptance of the E4 wristband as biosensor, are evaluated in this study. For the self-reported measures the perceived burden, compliance rates and craving variability of two types of EMA cue type measures were assessed; (1) interval-contingent measures and (2) signal-contingent measures [9]. In interval-contingent designs, a fixed number of measures are taken according to a standard schedule of intervals. With signal-contingent designs, a fixed number of measures are taken at randomly scheduled intervals, possibly within specified periods of time. Regarding the E4 wristband three factors were evaluated to determine the user acceptance of the device: (1) usability [10], (2) wearing comfort [11], and (3) perceived stigmatization [11] of the wearable.

1.1 Perceived Burden of the EMA Design

The perceived burden, variability and compliance of self-reported craving will be explored to determine which cue type (interval or signal) is most feasible for a monitoring study.

Interval-Contingent Design. The strength of interval-contingent is that the measures can be taken at certain times of the day at which it is likely that construct of interest will occur [12], in this case craving. A risk is that a at unexpected moments at which craving could occur could be missed if the interval time is not selected properly, which causes a systematic bias [13]. Another weakness is that an interval-contingent design can result in high levels of predictability: patients can predict the timing of interval assessments. This may alter their behavior in preparation of the recording time [13].

Signal-Contingent Design. The main advantage of a signal-contingent measures is that the occurrence of craving during multiple different time frames of the day can be examined [7]. A weakness is that signal-contingent recordings can sometimes be perceived obtrusive and disruptive or a recording can be missed because the signal was unexpected or could not be completed at the moment of the signal [13].

Compliance Rate EMA. Compliance rates are associated with the persuasiveness of a system, since the system will be used more when it is highly persuasive [1]. Compliance rates in EMA studies vary from 90% to 50%, but many studies have compliance rate around 75–80% [6]. It is not clear what causes this variation, but the contingent design is likely to affect these compliance rates.

Variability EMA. Researchers of EMA studies found low levels of self-reported craving in alcoholics [6]. It is not clear whether this is a genuine finding or that alcoholics are unwilling to admit their craving. This pilot study explores the variability of self-reported in two contingent designs.

1.2 User Acceptance E4

To determine user acceptance of the E4, usability, wearing comfort and perceived stigmatization are tested in the pilot. The E4 has the size of a large watch and is designed for continuous, real-time data acquisition in daily life.

Usability E4. It is important that the wearable biosensor has high usability, so that the participants use the device as intended by the researcher. Usability is described as the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment [10].

Wearing Comfort E4. An important aspect in integrating and accepting the E4 sensor in the respondents daily life is the wearing comfort. Fensli and Boisen [11] found several factors that play a role in this perceived comfort of wearable technology, for example perceived burden during daily activities like bathing or sleeping. Users preferred small discrete sensors that are as much as possible integrated in everyday objects [14] and did not affect daily behavior. Furthermore, they highly valued a comfortable, compact, reliable and easy to operate device. However, the E4 has some specifications that can not be adjusted, such as the device has to be charged every few days, cannot be worn during showering or extreme rain and the size and comfort of the device cannot be modified. It is explored to what extent this influences the user acceptance of the wearable.

Stigmatization E4. Fensli en Boisen [11] also found that people can experience stigmatization when wearing a sensor. Bergmann, Chandaria and McGregor [14] showed that users prefer sensors that cannot be seen by other people. This preference could be even more important for alcoholics, who often want to keep their addiction hidden from their social environment, including family and friends [15]. Stigmatization might cause respondents not to wear the wristband at certain moments (for example at work) or stop wearing the wristband altogether, posing a problem with compliance.

2 Methods and Materials

This study was a mixed methods feasibility pilot, in which both physiological and self-reported EMA data were collected. To explore reasons for possible (non)compliance, multiple questionnaires and an exit interview regarding perceived burden and user acceptance were performed.

2.1 Participants

In total, eight participants (3 male) between 19 and 24 years (M = 21.5, SD = 1.77) took part in the experiment over a period of eight days. Four (2 male) of these participants completed the whole experiment, the other four were only included for a part of the study, namely wearing comfort and stigmatization, since they did not perform the EMA part of the study. Participants were included if they met at least 2 of the 11 diagnostic criteria of the DSM-V for an alcohol use disorder, if they drank more than 14 glasses on average of alcohol per week and did not have a diagnosis of dependence of another substance than alcohol (except nicotine). Participants also had to own a smartphone and have access to a laptop or computer. The participants were college students from the University of Twente and received course credits to participate in the study. The Ethical Commission of the University of Twente approved this study and participants signed an informed consent before entering into the study.

2.2 Materials

EMA Recording. For recording the daily questionnaires a smartphone app named 'UTSurvey' [16] was used. The app issues an notification when a questionnaire had to be administered.

Physiological Data. For uploading the data of the E4 wristband, the program 'Empatica Manager' was used.

2.3 Procedure

Data-acquisition started with a meeting during which informed consent was obtained and the DSM-V criteria were assessed. Participants were then trained to use the EMA app, use the E4 sensor, charge the wristband and upload the data. During the week, the participants had to wear the E4 during the day, and charge it at night when they were sleeping. Additionally, they had to upload the recorded data at least every two days. In the EMA app, participants had to fill in an alcohol registry every morning, and complete a brief questionnaire four times a day. The first four days, phase one of the experiment, an interval-contingent design was used for the EMA app. The time slots at which questions had to be administered were predefined and the same for every day (11 am, 3 pm, 7 pm and 12 midnight). This final time slot was not mandatory since some participants might already be asleep. The second set of four days, the second phase, a signal-contingent design was employed. The time slots were randomized; between 9:00 am and 12:45 pm, between 1:00 pm and 4:45 pm and between 5:00 pm and 9:00 pm. The final time slot of the day was again at midnight, again not mandatory. After the experiment the participants administered online questionnaire about the perceived burden of the EMA designs, a questionnaire with the 'System Usability Scale' (SUS) about the usability of the E4 wristband and an exit interview on the wearing comfort and the perceived stigmatization of the E4 wristband.

2.4 EMA Measures

Morning Report. Upon awakening, the app questioned the participants about the time and number of standard drinks consumed on the previous day.

Prompts. Participants had to answer three questions after a prompt: alcohol craving, mood and coping ability. Since craving was the only construct of interest for this pilot and the two other questions were merely included to represent a realistic EMA burden, only craving will be further explained. Craving for alcohol was measured on a 10-point Likert Scale ranging from 1 (not at all) to 10 (very much). A single item measure of craving is a straightforward and time effective manner for assessing the level of subjective craving of a participant [17].

2.5 Measures

Perceived Burden of the EMA Design. For measuring the perceived burden of the participants for the EMA designs, an online open ended questionnaire was taken (See Table 1.) (translated from the original Dutch items).

Table 1. 11 Questions on perceived burden of the two EMA design

1.	What did you think of answering a questionnaire every day? Please explain.
2.	Would you be able to integrate answering (daily) questionnaires in your day-to-day life? Why could or couldn't you?
3.	What is your attitude towards daily surveys? Please explain.
4.	How did the continuous answering of questions go? Please explain.
5.	How difficult was answering the questions for you? Please explain.
6.	How burdensome did you find it to answer questions every day? Please explain.
7.	Did anything irritate you while answering the questions? If yes, what and why?
8.	Did anything go wrong with answering the questions? If yes, what and why?
9.	What did you think of answering questions on an app? Please explain.
10.	Did you think the app was easy to use? Please explain.
11.	Did you miss many administration moments? If yes, why?

Usability E4. Usability of the E4 wristband was measured with the System Usability Scale [10] consisting of ten 5-point Likert scale items. The SUS gives a global view of subjective assessments of usability [10].

Wearing Comfort E4. A semi-structured interview was performed, starting with a non-directive question "How did you experience last week?" and becoming more and more directive, leading to the closing question "How would you describe wearing the sensor", when a participant did not give any information on wearing comfort in earlier questions.

Stigmatization E4. A semi-structured interview was performed with the same non-directive to directive structure. The opening question was "Did you have any conversations about your participation in the experiment in the last week?" leading to "What was the impact of the sensor on your feeling of anonymity?".

2.6 Data Analysis

EMA Design

Compliance Rate and Variability EMA. The EMA data was extracted from the program 'Limesurvey' [18]. The compliance rate and variability in craving for the self-reported measures per EMA design were determined. Because of the exploratory nature and small sample in this research no statistical test was performed. Instead, multiple graphs for every participant per design were made with craving measurements for every time slot in order to make the variability and distribution of the craving scores visible.

Perceived Burden EMA. The questionnaire for the perceived burden was analyzed by filtering out the main statements relevant to the perceived burden of participants. After which the statements were labeled. The labeling continued until no new labels could be given. The inter-rater reliability was determined using Cohens Kappa, by labeling all data by a different researcher.

User Acceptance E4 Wristband

Compliance Rate E4. The physiological data was extracted from the Empatica Manager. The times on which data was recorded were visualized in a graph in order to evaluate the compliance rates.

Usability E4. The SUS scores were calculated according to [10]. With the use of the adjective rating scale [19], the SUS score was interpreted.

Wearing Comfort and Stigmatization E4. Wearing comfort and stigmatization were measured using a semi-structured interview. The text was divided in relevant fragments and labeled, as close as possible to the text of the fragment. The labeling continued until no new labels could be given. The inter-rater reliability was determined using Cohens Kappa, by labeling 12.5% of the data by a different researcher.

3 Results

The female participants reported drinking an average of 12.8 glasses of alcohol a week (sd = 6.53 glasses) divided over four evenings during the week (sd = 1.45 days). The male participants reported drinking 14 glasses (sd = 3.46 glasses) on average divided over three evenings during the week (sd = 0.69 days).

3.1 Self-reported EMA data

This research explored the variability and the compliance of alcohol craving measurements when using interval- and signal-contingent designs. During one time slot no prompt occurred and during another it was not possible for the participants to administer the questions due to technical problems. However, these minor anomalies are not expected to have influenced the results, as they constituted two out of the total of 81 completed time slots.

Compliance Rate and Variability EMA. Participants reported craving greater than 1 relatively often: out of 81 completed time slots, noticeable urge to drink was reported 42 times. The average craving was 3.68 (SD = 3.36) with the interval-contingent design and 3.05 (SD = 2.90) with the signal-contingent design. Figure 1 visualizes the compliance and variability of alcohol craving per participant. It seems that the compliance and variability of the craving measurements between the interval- and signal-contingent design do not differ greatly.

The average compliance in the interval-contingent design was 67% and in the signal-contingent design 61%. The craving scores were higher in the evening compared to the morning and afternoon. As expected, the number of completed time slots was much lower at the non-mandatory time slot at midnight (time slot 4, 8, 12 & 16). Participant four mentioned she missed a lot of time-slots, due to not having an Internet connection or beeing occupied at work.

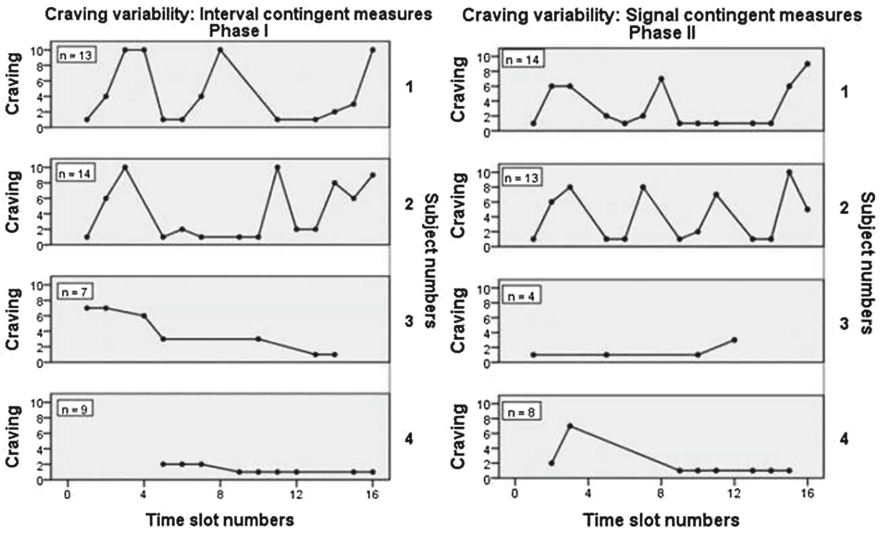

Fig. 1. Craving scores per participant over time. The missing values are left out (n = the number of administered data points). The time slot numbers are displayed on the horizontal axes. Every four numbers represent one day: First time slot = morning, second time slot = afternoon, third and fourth time slots are evening.

Perceived Burden EMA. The statements about the perceived burden had an inter-rater reliability of 0.71 which corresponds with a substantial strength of agreement. The effort to complete the questions at a single time slot was generally experienced as low. All participants thought the questions were easy to answer: "Answering the questions went well, the questions were easy and you could answer in a detailed, but easy manner, how you felt at the moment." (Participant 4, q.4). The questionnaire at midnight was sometimes perceived as an higher burden: "(...) the questionnaire at midnight was sometimes unpractical (I was at the cafe without my phone then or was sleeping already)." (Subject 2, q.1). The burden of filling in the questionnaires with a frequency of four times a day was perceived as relatively high by two participants. They reported feeling bad or stressed because of the questionnaires: "I don't like the feeling of always having to answer something. I felt very pressured. When I was not able to answer questions I felt bad because I did not finish the assignment" (Participant 4, q.6). However, others mentioned that they did not feel it as a burden to answer the questions multiple times a day: "Not really a burden, you have your phone often near you anyway." (Participant 1, q.6). The burden was heightened with the signal-contingent design, compared to the interval-contingent design: "It is easier if the time slots are at a set time, in that manner you can anticipate to react and after a while you don't forget to fill it in anymore." (Participant 1, q.2).

3.2 User Acceptance E4 Wristband

Compliance Rate E4. The compliance rate of wearing the E4 differed over participants. An overview of the times each participant wore the E4 can be seen in Fig. 2. Note that participant four only wore the E4 on 4 of the 7 days and participant three wore the E4 every day, however sometimes only for a few hours (see for example Monday). The other two participants all had a higher compliance.

Usability E4. The usability of the E4 wristband when wearing it on a daily basis for a week, with use of the System Usability Scale was on average 65 points on a 100-point scale, which is considered to be slightly below average [10]. Scores on individual questions could range between 1 and 4. The participants thought the E4 was easy to use (3.25) and did not feel like they needed a lot of technical support (1). However, all but participant 4 did not really wanted to use the E4 frequently (1.25) and found the E4 sometimes frustrating (3).

Wearing Comfort E4. The agreement rate regarding the qualitative analysis had a Cohens kappa of 0.77 which corresponds with a substantial strength of agreement. The general impression of the participants on wearing the sensor was neutral to positive and was described by all participants as "fine". All participants indicated that they had to get used to the sensor, but this happened quickly and with ease. Participant one remarked that it was fine to wear

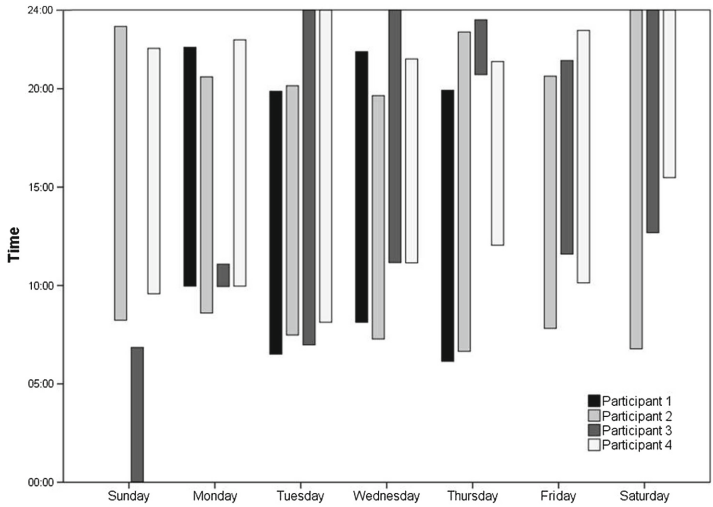

Fig. 2. Time each participant wore the E4 over a course of a week.

the sensor for a week, but would not have wanted to wear the wristband any longer. The participants mentioned that it is burdensome that the sensor is only splash-waterproof. For many activities, like doing the dishes, showering or even cycling through the rain, it was not possible to wear the wristband. One participant indicated that he found it annoying to have to remember to bring the charger, charge the wristband and upload the data. Multiple participants did mention that the sensor sometimes pressed painfully on the bone and was either too tight or too loose. One of the men, as well as both females found the sensor "ugly" and considered it to be too big to be worn with tight clothes. A final negative point about the comfort that was mentioned by the participants was that the sensor is not easy to wear during exercise, since in some sports wearing anything around your wrists is prohibited and some are too "rough" for gathering good data.

Stigmatization E4. All participants indicated that they were approached by many people about the sensor. The majority of the participants therefore became more aware of the sensor. Especially family, friends and acquaintances addressed the participants about the sensor. All participants indicated that they had a positive feeling about the conversations and explaining why they had to wear the sensor. One participant mentioned that if he should wear the sensor for longer, it would become irritating that his environment would continue to question the sensor. The participants indicated that the sensor was often thought of as a watch or sports and fitness tracking device. Three participants also said they were asked if they were wearing a house arrest curfew band due to a criminal offense.

4 Conclusion and Discussion

In order to make more informed design decisions when carrying out a full-scale monitoring study in the field of alcohol craving, the current study pilot tested two EMA designs and the validity of both physiological and self-report measures. The current study showed that both the interval- and signal-contingent EMA study designs can capture relevant and meaningful variability of self-reported alcohol craving. This is in contrast with prior mentioned findings [6], however students might be more willingly to admit their craving then alcoholics. Therefore, it is not definite that these results will transfer to the target population, but a relevant finding is that no significant difference between the two designs was found. The overall compliance of the self-reported measures was lower (i.e. 64%) then found by multiple EMA studies [6], but highly different among participants. The type of EMA design did not seem to explain the differences in compliance rates. Further research should focus on explaining the variations in compliance rates.

The burden of filling in multiple questionnaires a day was perceived as quite high, the interval-contingent was preffered over the signal-contingent design, as it was perceived as less obtrusive. It should be noted, though, that this preference for interval-contingent EMA may also be explained by an order effect, as all participants started with this design and the perceived burden may have increased in course of time. Nevertheless, all participants experienced both designs, allowing a direct comparison of perceived burden within subjects.

All the participants found the E4 easy to use, rated the usability slightly below average and described the wristband as fine. Women experienced less wearing comfort, due to the size of the E4. Participants found it annoying that the wristband is only splash water proof. Despite the positive rating of the E4 wristband, the compliance rate of the physiological data was quite low for two of the participants. As mentioned, addicts are notoriously non-compliant to schedules and instructions [6]. An effort should be made to increase this compliance, we therefore propose for the monitoring study to give participants a micro incentive to stimulate compliance to both the self-reported and physiological data. Musthag, Raij, Ganesan, Kumar, and Shiffman [20] showed these incentive studies to be low cost, while ensuring high compliance, good data quality, and lower retention issues.

A potential threat to the user acceptance of the E4 may be the apparent visibility and the attracted attention of the wearable biosensor, which may particularly apply to the target population of the monitoring study, persons with addiction, might not prefer this scrutiny, and consider this to be a stigmatization. Unfortunately no other research device is available that unobtrusively measures heart rate and electrodermal activity. Therefore, we propose to prepare the participants by providing coping strategies. For example, the participant can learn to respond that the sensor is helping to improve their health. It should be explored whether this lowers feelings of stigmatization.

In conclusion, the present feasibility study showed that an interval-contingent EMA design with micro incentives and the E4 wristband are feasible and low

burden design decisions for a real time, alcohol craving study. This finding has relevance for similar studies on other addictive behaviors, but may also contribute to less related field of research (e.g. occupational stress). For the latter types of studies the present work can also be seen as a blueprint on how to do a thorough, relatively quick formative evaluation for a subsequent daily life study. Daily life studies are often rather time consuming, costly and put a heavy burden on participants [3]. It is therefore essential to make well substantiated choices for the study design based on both literature and a feasibility pilot, such as the one presented here.

References

1. Oinas-kukkonen, H., Harjumaa, M.: Key issues, process model, and system features persuasive systems design: key issues, process model, and system features. Commun. Assoc. Inf. Syst. **24**(28), 485–500 (2009)
2. Rohsenow, D.J., Monti, P.M.: Does urge to drink predict relapse after treatment? Alcohol Res. Health **23**(3), 225–232 (1999). The Journal of the National Institute on Alcohol Abuse and Alcoholism
3. Csikszentmihalyi, M., Mehl, M.R., Conner, T.S.: Handbook of Research Methods for Studying Daily Life. Guilford Publications, New York (2013)
4. Carter, B.L., Tiffany, S.T.: Meta-analysis of cue-reactivity in addiction research. Addiction **94**(3), 327–340 (1999)
5. Takarangi, M.K.T., Garry, M., Loftus, E.F.: Dear diary, is plastic better than paper? I can't remember: comment on Green, Rafaeli, Bolger, Shrout, and Reis. Psychol. Methods **11**(1), 119–122 (2006). Discussion 123–125
6. Shiffman, S.: Ecological momentary assessment (EMA) in studies of substance use. Psychol. Assess. **21**(4), 486–497 (2009)
7. Hufford, M.R., Shiffman, S.: Methodological issues affecting the value of patient-reported outcomes data. Expert Rev. Pharmacoeconomics Outcomes Res. **2**(2), 119–128 (2002)
8. Scollon, C.N., Kim-prieto, C., Diener, E.: Experience sampling: promises and pitfalls, strengths and weaknesses. J. Happiness Stud. **2003**(4), 5–34 (1925)
9. Serre, F., Fatseas, M., Swendsen, J., Auriacombe, M.: Ecological momentary assessment in the investigation of craving and substance use in daily life: a systematic review. Elsevier Ireland Ltd. (2015)
10. Brooke, J.: System usability scale (SUS): a quick-and-dirty method of system evaluation user information. In: Usability Evaluation Industry, vol. 189, pp. 4–7 (1996)
11. Fensli, R., Boisen, E.: Human factors affecting the patient's acceptance of wireless biomedical sensors. In: Fred, A., Filipe, J., Gamboa, H. (eds.) Biomedical Engineering Systems and Technologies, BIOSTEC 2008. Communications in Computer and Information Science, vol. 25, pp. 402–412. Springer, Heidelberg (2008)
12. Moskowitz, D.S., Russell, J.J., Sadikaj, G., Sutton, R.: Measuring people intensively. Can. Psychol. Psychol. Can. **50**(3), 131–140 (2009)
13. Smyth, J., Wonderlich, S., Crosby, R., Miltenberger, R., Mitchell, J., Rorty, M.: The use of ecological momentary assessment approaches in eating disorder research. Int. J. Eat. Disord. **30**, 83–95 (2001)
14. Bergmann, J.H.M., Chandaria, V., McGregor, A.: Wearable and implantable sensors: the patient's perspective. Sensors (Switzerland) **12**(12), 16695–16709 (2012)

15. Howard, M., McMillen, C., Nower, L., Elze, D., Edmond, T., Bricout, J.: Denial in addiction: toward an integrated stage and process modelqualitative findings. J. Psychoactive Drugs **34**(4), 371–382 (2002)
16. Trompetter, H.R., Borgonjen, F., Zwart, M., van Tongeren, S.: UTSurvey Manual v1 (2015). http://www.utwente.nl/igs/datalab
17. Drobes, D.J., Thomas, S.E.: Assessing craving for alcohol. Alcohol Res. Health **23**(3), 179–186 (1999). The Journal of the National Institute on Alcohol Abuse and Alcoholism
18. Schmitz, C.: Limesurvey: an open source survey tool (2012). http://www.limesurvey.org
19. Bangor, A., Kortum, P.T., Miller, J.T.: An empirical evaluation of the system usability scale. Int. J. Hum. Comput. Interac. **24**(6), 574–594 (2015)
20. Musthag, M., Raij, A., Ganesan, D., Kumar, S., Shiffman, S.: Exploring micro-incentive strategies for participant compensation in high-burden studies. In: Proceedings of the 13th International Conference on Ubiquitous Computing - UbiComp 2011, p. 435 (2011)