UNIVERSITY OF TWENTE.

Faculty of Behavioural, Management and Social Sciences (BMS)

A feasibility study into measuring intraindividual craving in a longitudinal study

Measuring self reported craving in a naturalistic setting combined with the usability of the Empatica E4 wristlet

Jessica Dijanne Stroes

Bachelor thesis

June 2016

Supervisors:
Dr. M.L. Noordzij
Dr. M.E. Pieterse
Msc. H.G. van Wijk

Cognitive Psychology and Ergonomics (CPE)
Faculty of behavioural, management and social sciences (BMS)
University of Twente
P.O. Box 217
7500 AE Enschede

Table of Contents

1. Abstract 3
2. Introduction 4
   2.1 Context 4
   2.1 Stakeholders 4
   2.2 Elaboration of the study 5
      2.2.1 Variability of craving measurements 5
      2.2.2 Perceived burden and compliance of IRM-NS designs 8
      2.2.3 Usability of the Empatica E4 wristlet 9
3. Method 9
   2.1 Design 9
   2.2 Participants 10
   2.3 Measures 10
   2.4 Material 12
   2.5 Procedure 12
   2.6 Data analysis 13
4. Results 14
   4.1 Variability of craving measurements 14
   4.2 Perceived burden and compliance of the IRM-NS designs 16
      4.2.1 Perceived burden of the IRM-NS designs 16
      4.2.2 Compliance of the IRM-NS designs 18
   4.3 Usability of the Empatica E4 wristlet 18
5. Discussion 19
   5.1 Exploring the variability of craving measurements 19
   5.2 Perceived burden and compliance of the IRM-NS designs 21
      5.2.1 Perceived burden of the IRM-NS designs 21
      5.2.2 Compliance of the IRM-NS designs 22
   5.3 Usability of the Empatica E4 wristlet 23
   5.4 Limitations of the study 24
   5.5 Recommendations for further research and conclusion 25
6. References 27
7. Appendix 31

1. Abstract

Introduction This study served as a feasibility study for a PhD monitoring study that will be done in cooperation with Tactus (an addiction treatment centre in the Netherlands). This PhD study will monitor ex alcohol addicts in their daily life and register their self reported psychological craving thoughts combined with measuring their physiological responses. The current study explored the variability of self reported craving for alcohol measurements in a small group of students with the use of 'Intensive Repeated Measures in Naturalistic Settings' (IRM-NS), specifically with the use of interval-contingent and signal-contingent designs. Furthermore it explored the perceived burden and compliance of the use of IRM-NS designs and the usability of the Empatica E4 Wristlet.

Method Four students with an above average alcohol usage reported with a frequency of four times a day the amount of craving they experienced combined with wearing the Empatica E4 wristlet, both for the period of one week. Afterwards questionnaires were filled in about the perceived burden of the design. The System Usability Scale was used to measure the usability of the E4 wristlet.

Results Both designs measured variability of self reported craving. The measures are similar to expectations (e.g. high craving in the evening, for students drink most in the evening (Kuntsche & Labhart, 2012). The burden of filling in the questionnaires with a frequency of four times a day as perceived by the participants was different per person, it ranged from neutral to negative. The compliance of participants was relatively low. The usability of the E4 wristlet was according to the adjective rating scale of Bangor, Kortum and Miller (2009) 'good'.

Conclusion Regarding the variability of craving, the use of an interval-contingent as well as an signal-contingent design seems suited. However, recommended is to use an interval-contingent design for the PhD monitoring study because participants can anticipate better on the questionnaires due to the set times. The E4 wristlet has a reasonable usability, it should be taken into account however that the wearability of the wristlet might be low.

2. Introduction

2.1 Context

Globally, alcohol use disorders are one of the most prevalent mental disorders, affecting an estimated 3.6% of the worldwide population between 15 and 64 years (Rehm, Shiedld, Gmel, Rehm, Frick, 2012). Alcohol use disorders bring a significant health burden with it, directly- as a disease and indirectly as a risk factor (Rehm et al., 2012). Miller et al. (as cited in Bottlender & Soyka, 2004) state that despite the effectiveness of treatment for alcohol use disorders, relapse is still common after treatment. Estimated rates of relapse after treatment vary widely but are typically in the range of 40 to 60% within the first few months after treatment and as high as 70 to 80% by the end of the first year after treatment (Dawson, Goldstein & Grand, 2007).

A variable which can be a predictor for relapse after treatment is craving, an important symptom in the DSM-5 classification for alcohol use disorders (AUD). Craving generally refers to 'the desire or urge to reexperience the effect of a previously experienced psychoactive substance' (UNDCP/WHO, 1992). In a study of Bottlender and Soyka (2004), patients who relapsed after treatment showed higher scores on the obsessive-compulsive drinking scale (OCDS) than patients who did not relapse. Comparable results were found in other studies (Evren, Durkaya, , Evren, Dalbudak & Cetin., 2012; Gordon et al., 2006). To prevent ex-addicts from having relapse it is important that predictors of relapse, such as craving, are recognized by patients. With a patient being aware of a predictor like craving, treatment for alcohol use disorders can be improved. Preventive techniques (see for example Stalcup, Christian, Stalcup, Brown & Galloway, 2006) can be put into action as soon as a patient becomes aware of craving.

2.1 Stakeholders

This study serves as a feasibility study for a PhD monitoring study that will be done in cooperation with Tactus (an addiction treatment centre in the Netherlands). This PhD study will monitor ex alcohol addicts in their daily life and register their self reported psychological craving thoughts combined with measuring their physiological responses such as their heart rate variability (HRV) and electro dermal responses (EDR). The final aim of the PhD project is to develop a remote monitoring and coaching system to give feedback to the client when craving responses are increased

in order to support clients with preventing a lapse. The system will give feedback based on physiological information from a wearable biosensor. This system is to be used as an additional feedback system for ex alcoholics who are treated with the existing Tactus treatment 'Alcoholdebaas.nl'. Requirements for the current feasibility study are that the study has to be conducted with within person research in a real time, real world setting.

2.2 Elaboration of the study

Different research statements are composed to examine the feasibility of certain aspects of the PhD study. Each research statement will be mentioned below and clarified with some background information.

2.2.1 Variability of craving measurements

To gain insight in the occurrence of craving of ex-addicts during the day in a real time and real world setting and because craving fluctuates rapidly (Tidey et al., 2008), multiple measures per day have to be conducted. 'Intensive Repeated Measures in Naturalistic Settings' (IRM-NS) are well suited for the assessment of within-person variability (Moskowitz, Russell, Sadikaj and Sutton, 2009). There are three primary designs for the collection of self reports in IRM-NSs: (1) time-contingent designs; (2) signal-contingent designs and (3) event-contingent designs. In time-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 signals, possibly within specified periods of times. With event-contingent designs, participants report if a defined event occurs (Moskowitz et al., 2009).

Compared to retrospective designs (where measurements are done after the event has taken place), IRM-NS designs have the advantage that they are less susceptible for memory biases. Errors in cognitive processing and memory errors decrease when participants reflect on short periods of time that are very recent (Schwarz, 2012). In retrospective designs, responses tend to be overly influenced by current mood state, the most recent and salient experiences and believes about the self (Bolger, Davis and Rafaeli, 2003). Another advantage of IRM-NS designs is that it allows researchers to study how people experience alcohol in an ethically sound way.

Laboratory studies of the mechanisms of heavy drinking are often restricted by legal and ethical concerns. By using an IRM-NS design, researchers can study these mechanisms with little restrictions about the actual drinking of alcohol by the subjects for instance (Wray, Merrill & Monti, 2014).

A drawback of using a repeated intensive assessment is the possibility that the higher burden for the participant compared to a retrospective design discourages participation. This can result in a sampling bias where only certain participants complete the assessment which is related to motivation and personality factors (Scollon, Kim-Prieto & Diener, 2003). Another menace is measurement reactivity, this is when participants change their behaviour or responses because of the repeated measurements (French & Sutton, 2010). When participants repeatedly have to report the amount of alcohol they consume for example, they can change their alcohol consumption because they are more aware of their possibly harmful behaviour.

The three primary designs have different strengths and weaknesses. The main strength of event-contingent recording is that events are not to be missed because participants themselves initiate when to record them (Hufford & Shiffman, 2002). The recording does not rely on the chance of a signal taking place to sample the behaviour of interest (Smyth, 2000). The primary limitation is that event contingent recording completely relies on the compliance and competence of the participants. Participants must be willing and able to initiate the recording and to recognize craving in such a way that it strokes with the perception of the researchers. (Smyth, 2000). Research finds however, that with event-contingent recording a relative high risk for a low compliance rating for measuring craving is present (Litt, Cooney and Morse, 1998; Flannery et al., 2001). Furthermore, event-contingent recording requires high vigilance to have a high compliance with the real-time reporting, because respondents have to take action to report craving without a reminder. It also requires high vigilance to constantly monitor the specific experience being measured (Hufford & Shiffman, 2002).

With time-contingent measures, a strength is that the measures can be taken at certain times of the day at which it is likely that craving will occur. This prevents that times at which it is likely that craving occurs, may be undersampled (Moskowitz et al., 2009). A risk is that a systematic bias is introduced by assessing behaviour at fixed points in time (Wheeler & Reis, 1991). Unexpected moments at which craving could occur could be missed if the interval time is not selected properly, which causes

a systematic bias (Smyth et al., 2000). 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 behaviour in preparation of the recording time, which reduces the degree to which the data reflects the natural behaviour of the patient (Smyth et al., 2000).

The main advantage of signal-contingent design is the dynamic sampling that can be applied (Hufford & Shiffman, 2002). Because of this the occurrence of craving during the multiple time frames of the day can be examined. With a signal-contingent design, the possibility of a systematic bias that might be introduced by assessing craving at fixed points in time is eliminated because of the random times of the recordings (Wheeler & Reis, 1991). Also due to the random schedule of the recordings the participants cannot predict the recordings (Smyth et al., 2000). The primary limitation of a signal-contingent design is that the chance that the signal and the event will coincide decreases with the event in question being rarer (Wheeler & Reis, 1991, Smyth., 2000). Another weakness is that signal-contingent recordings can be obtrusive and disruptive sometimes, for example whenever a recording is required while being at a party (Wheeler & Reis, 1991). Data will be lost sometimes because the subject cannot complete the recording at the time of the signal (Smyth et al., 2000).

Because of the low compliance with event-contingent recordings in earlier studies measuring craving (Litt et al., 1998; Flannery et al., 2001) and the relative high vigilance needed to constantly monitor for specific experience of craving compared to the other two primary designs (Hufford & Shiffman, 2002), this form of recording is disregarded. It has been decided to explore the use of interval-contingent and signal-contingent recordings for the measurement of craving. Specifically the variability in craving measurements will be explored to discover if the designs are feasible for measuring self reported craving. Therefore the first and main research statement is as follows:

> *'Exploring the variability of measurements of alcohol craving when using interval-contingent recordings and signal-contingent recordings within an IRM-NS design.'*

2.2.2 Perceived burden and compliance of IRM-NS designs

A drawback of using an IRM-NS design is the burdensome or intrusive nature of the method (Scollon, Kim-Prieto & Diener, 2003). Although the recording moments require often little time to complete, subjects are nevertheless interrupted on several moments during the day and their attention is distracted from current activities (Alliger & Williams, 1993). The burden that an IRM-NS design places on participants can result in high attrition rates and noncompliance (Moskowitz et al., 2009). Signal-contingent design is the most burdensome for participants because of the randomness and frequency of the design (Conner & Bliss-Moreau, 2005). Interval-contingent designs are less burdensome because recordings are made at standardized times. In this way participants can adjust their schedules around the scheduled recording moments (Conner & Bliss-Moreau, 2005).

The burdensome nature of an IRM-NS method can be adjusted by changing the amount of recordings per day and the length of the recordings to be completed. According to Scollon, Kim-Prieto & Diener (2003) a general rule of thumb is: the more signals per day, the shorter the form (in this case recording) should be. By exploring the burden of completing a questionnaire multiple times a day as it is perceived by the participants themselves, possible recommendations can be given regarding the length and frequency of the recordings. By also comparing the two designs among each other, a possible recommendation can be done for the preferred use of one of the designs. Therefore, the second research statement is as follows:

> *'Exploring and comparing the perceived burden of subjects when completing a questionnaire about alcohol craving multiple times a day for a week when using interval-contingent recordings and signal-contingent recordings within an IRM-NS design.'*

The success of an IRM-NS study depends upon the compliance of participants (Christensen, Barrett, Bliss-Moreau, Lebo & Kaschub, 2003). Therefore the compliance of the interval- and signal-contingent designs will be explored as well as compared among each other. The third research statement is as follows:

> *'Exploring and comparing the compliance of subjects when completing a questionnaire about alcohol craving multiple times a day for a week when using interval-contingent recordings and signal-contingent recordings within an IRM-NS design.'*

2.2.3 Usability of the Empatica E4 wristlet

For measuring the physiological responses of the ex alcohol addicts a wearable bio sensor is used. It is important that the wearable bio sensor is easy to use, so that the participants can properly use the device as intended by the researcher. A term which can be used for this ease of use or user friendliness is 'usability'. Usability is described as “the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in a particular environment” (Brooke, Bevan, Brigham, Harker and Youmans, 1990). By measuring the usability of the Empatica E4, a possible recommendation can be done regarding the use of the Empatica E4 for further research. A popular scale for the measurement of the usability of a system or product is the 'System Usability Scale' (Brooke, 1996). This survey scale allows the researchers to quickly assess the usability of a product (Bangor, Kortum and Miller, 2008). The scale will be used to measure the usability of the Empatica E4 Wristlet roughly when users wear it for a week. The fourth research statement is as follows:

> *'Measuring the usability of the Empatica E4 wristlet with the System Usability Scale when daily wearing it for a week when taking part in a study about alcohol craving.'*

3. Method

This study is a feasibility study for a PhD study in progress at the University of Twente about real time telemonitoring of physiological processes and self-reported craving to reduce relapse rates in alcohol addiction. The method described below is conducted in cooperation with two other students from the University of Twente. Each student used different data from the same research conducted. Parts of the method that are not relevant for the current research are nevertheless mentioned for the method to be complete. For the execution of this research permission was asked and granted by the ethics committee of the faculty Behavioural, Management and Social Sciences (BMS) of the University of Twente.

2.1 Design

This is a qualitative longitudinal study that incorporates within subject designs consisting of 'Intensive Repeated Measures in Naturalistic Settings' (IRM-NS). The study was conducted in a real time, real world setting.

2.2 Participants

In total four (female) participants volunteered to participate with the study (M_{age} = 22.25, SD_{age} = 1.71). Participants were included if they met at least 2 of the 11 diagnostic criteria of the DSM-5 for an alcohol use disorder, if they drank more than 14 glasses on average of alcohol per week (Jellinek, n.d) and did not have a diagnosis of dependence of another substance than alcohol (except nicotine). The minimum age of the participants had to be 18 years old, in accordance with the regulations of serving liquor following the Dutch Alcohol Licensing Law. The participants were sampled by convenience sampling from the direct environment of the researchers. All the participants fitted the requirement of fluently speaking and understanding the Dutch language. The participants were required to be doing a study at an (applied) university and had to own a smart phone with an constant internet connection and access to a laptop or computer.

2.3 Measures

Two types of IRM-NS were used for presenting the tasks: interval-contingent measures and signal-contingent measures. For measuring the variability of craving of the two designs, half of the experiment was executed with an interval-contingent design and the other half with a signal-contingent design. The measurements were divided across the day: one for the morning and one for the afternoon and for the evening two. Of these four moments, three were used to vary between interval-contingent and signal-contingent measures. The fourth and last task of the day was carried out on a daily basis for both interval- and signal-contingent design at a predefined time, namely at midnight. The fourth time moment was not varied because in the part of the study with signal-contingent time slots, the last possible moment of the third interval was 9:00 pm. This would make the possible time frame for a fourth moment too small to vary within in order for it to have a time frame of four hours: it would only be three then because expected is that most participants will be sleeping after midnight.

The first four days, phase one of the actual experiment, interval-contingent measures were taken for all participants. The time slots at which the task had to be executed were pre defined and were the same for every day during the first phase. Participants were well informed about the exact time slots of the task so that they

were well aware of what to expect. The morning time slot was set at 11:00 am, the afternoon time slot at 3:00 pm and the evening time slot at 7:00 pm. The final time slot was set at 12:00 midnight but was not mandatory since some participants might be asleep already.

The second set of four days, the second phase, a signal-contingent design was used for all participants. The time slots were randomized within the pre defined time frames. The participants were not informed about the exact time slots. However, they were informed about the time frames that the time slots would fall in. The time frame of the morning was between 9:00 am and 12:45 pm, the time frame of the afternoon was between 1:00 pm and 4:45 pm and the time frame of the evening was between 5:00 pm and 9:00 pm. The final time slot of the day was at midnight. In the instructions it was mentioned that it does not matter if participants miss the final time slot of 12:00 if they would already be asleep. For the randomized time slots used in this study see appendix A.

Every moment of measurement, the same task was executed: three identical questions were asked (Appendix B). The task in the morning had an extra question about the alcohol usage the previous day. A time of two minutes was estimated to conduct the total task. The questions were asked using an app on the smartphone of the participants. There was chosen to present the questions on a smartphone because a smartphone can prompt participants when a task has to be completed. This increases the compliance of completing the tasks but also may reduce the perceived burden **(**Takarangi, Garry and Loftus, 2006). Each task was available to fill in for 1,5 hour. Of the four items in the task, one item was important for the aim of this research: 'Hoeveel trek in alcohol heb je op dit moment? (How much craving do you have for alcohol at this moment?). The other three questions were used for another research in preparation and to simulate a questionnaire size which is more realistic for future research so that a realistic perceived burden of the participants was measured. The craving for alcohol was recorded on an 10-point Visual Analogue Scale (VAS) 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 (Drobes & Thomas, 1999), though it is limited in the ability to provide information about multiple elements that can define the experience of craving (Tiffany, 1992). In the light of this research it is only important to assess if there is craving at all without assessing the underlying elements, so a one item measure of

craving will suffice. Furthermore it is important that the burden of completing the questionnaire multiple times a day as low as possible.

For measuring the perceived burden of the participants for completing the questionnaire multiple times a day, an online questionnaire was executed at the end of the experiment, consisting of 11 self-made items (see appendix D).

For measuring the usability of the E4 wristlet the System Usability Scale (Brooke, 1996) was conducted. The SUS is an ten-item scale with forced choice statements where the respondent indicates the degree of agreement or disagreement with a 5-point Likert scale. The SUS gives a global view of subjective assessments of usability (Brooke, 1996) and can be found in Appendix E.

2.4 Material

For recording the daily questionnaires an app named 'UTSurvey' (made by Frank Borgonjen) was used. Participants could download this app on their smartphone and the app gave an alert when a questionnaire had to be filled in.

For measuring real-time physiological measures, a wearable device named the 'E4 Wristlet' (Empatica, model E4 - 2015) was used, including accompanying accessorises (an USB cable and docking station for the wristlet). For uploading the data of the E4 Wristlet, the program 'Empatica Manager' was used. The laptop owned by the concerning participant was used by the participants for uploading data from the E4 Wristlet.

2.5 Procedure

The day before the IRM-NS recordings the participants had to fill in an informed consent (appendix C), a questionnaire about their alcohol use and the DSM-5 for an alcohol use disorder (appendix F,G). On the same day a instruction was given to the participants about the procedure of the research. Specifically about the daily questionnaire and the E4 wristlet. The apps and programs needed for the research were downloaded by the participants after reading the instruction sheets. The researchers were present to answer questions and to help with downloading the programs necessary. The instruction documents that the participants received can be found in appendix H. During the next eight days the participants had to conduct a short questionnaire with a frequency of four times daily. The participants were prompted by the researcher at certain set times when they had to complete the task.

After the experiment the participants received an online questionnaire about the perceived burden of filling in the questionnaires and the SUS about the usability of the E4 wristlet (Appendix D, E).

During the eight days the participants had to fill in the questionnaire daily they also had to wear the E4 wristlet the whole day, except when sleeping. Every day the E4 wristlet had to be charged and every two days the data from the E4 wristlet needed to be uploaded by the participants. Because the E4 wristlet was not waterproof, participants had to remove it when showering or when doing other activities were water was involved.

2.6 Data analysis

All data analyses was done with SPPS statistics (version 23). The data was extracted from the program 'Limesurvey'. The mean and standard deviation of the craving scores were calculated per person, design and time frame and also for both designs together. The compliance rate per participant, per design and an average compliance rate overall was calculated. Furthermore, multiple line graphs and box plots for every participant per design were made with the collected craving measurements for every timeslot in order to make the variability and distribution of the craving scores visible. The questionnaire for the perceived burden was analysed by the researcher itself, who filtered out the main statements relevant to the perceived burden of participants. The inter-rater reliability was determined with help of one person who received an explanation about what is meant by the perceived burden. This person also filtered out the main statements relevant to the perceived burden of the participants and these answers were compared with that of the researcher, from which Cohen's kappa was calculated. Cohen's kappa was interpreted with help of benchmarks created by Landis and Koch (1977). The SUS scores were calculated according to the article of Brooke (1996). With the use of the adjective rating scale the SUS score was interpreted (Bangor et al., 2009).

4. Results

The main goal of this research was exploring the variability of alcohol craving measurements when using time-contingent and signal-contingent measures. Because of the explorative nature of this research no statistical analysis was used. With one timeslot it was impossible for the participants to execute the tasks since the questionnaire did not work due to technical problems. Also one timeslot the researcher failed to send a text message to prompt the participants. These minor anomalies are not expected to have much influence on the results. Below in table 1, an overview of participant information can be found.

Table 1

Participant information

Participant ID	% Compliance IC design	% Compliance SC design	DSM-5 score	Average alcohol units p/w	SUS score
1	81.25	87.50	5	18-22	67.5
2	87.50	81.25	6	>24	50
3	43.75	25.00	0	12-18	62.5
4	56.25	50.00	5	12-18	80

Note. A score of 2-3 on the DSM-5 for an alcohol use disorder indicates a mild AUD. The SUS score can range from 0-100

4.1 Variability of craving measurements

Participants reported some craving (i.e. urge greater than 1) relatively often: out of 81 completed time slots, noticeable urge to drink was reported 42 times. The average urge to drink was 3.68 ($SD_{intervalurge}$ = 3.36) with the interval-contingent design and 3.05 ($SD_{signalurge}$ = 2.90) with the signal-contingent design, for both designs together the average urge was 3.39 (SD_{urge}=3.15). It should be taken into account that one (participant nr. 3) of the four participants had a much lower DMS-5 score than the other participants, this might have had effect on the amount of craving experienced during the experiment.

In Table 2 an overview is made of the craving scores per timeframe. Noticeable is that the average craving score for both designs together is higher for the

last two time slots in the evening compared to the first two in the morning and afternoon. The number of completed timeslots is much lower for timeslot 4.

Table 2

Average craving per time frame

Timeframe	Average craving interval- and signal-contingent recordings	Number of completed time slots in total
1: 09:00 am - 12:45 pm	1.46 (SD = 1.29)	n = 24
2: 01:00 pm - 04:45 pm	2.81 (SD = 2.46)	n = 16
3: 05:00 pm - 09:00 pm	4.86 (SD = 3.52)	n = 22
4: 12:00 midnight	6.17 (SD = 3.59)	n = 12

Note. The average craving can range from 1-10. The total possible amount of time slots per timeframe was 32

To demonstrate the within person variability, a box plot per design was made for each participant as seen in Figure 1. The distribution of the scores differs per subject: The first and second person show a high variability within the scale whereas the third and fourth person show a low variability within the scale.

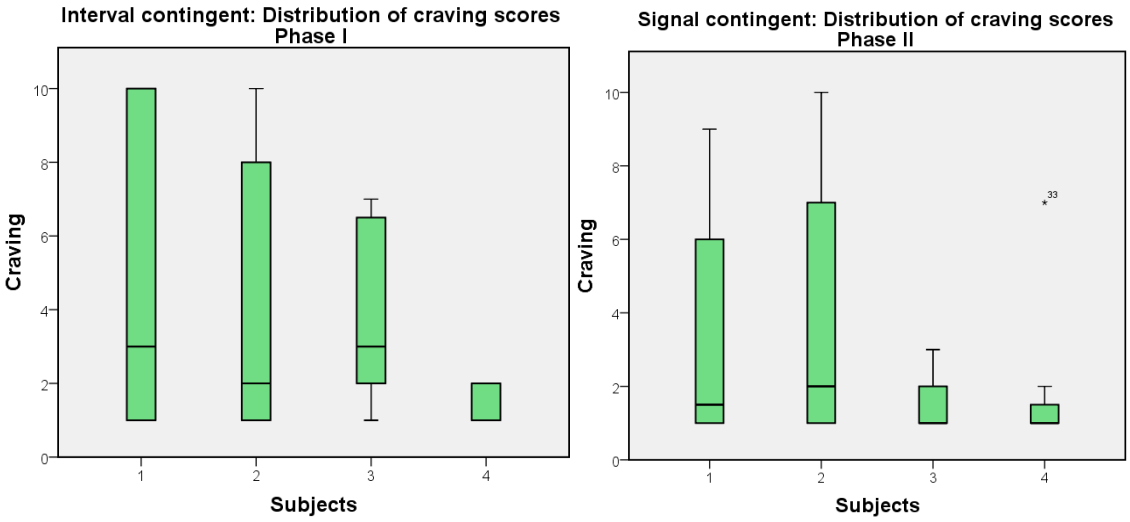

Figure 1. Distribution of craving (range 1-10) per person per design

Figure 2 visualizes the variability of alcohol craving per subject as measured with an interval-contingent and a signal-contingent design, in a different manner. As visible in

Figure 2, all craving scores are measured at least once. Initially it seems that the variability of the craving measurements between the interval- and signal contingent design do not differ greatly. It is noticeable that every subject seems to have some sort of trend (e.g. high variability in craving scores, stable scores of craving) regarding the scoring of craving which continues in both measurement designs.

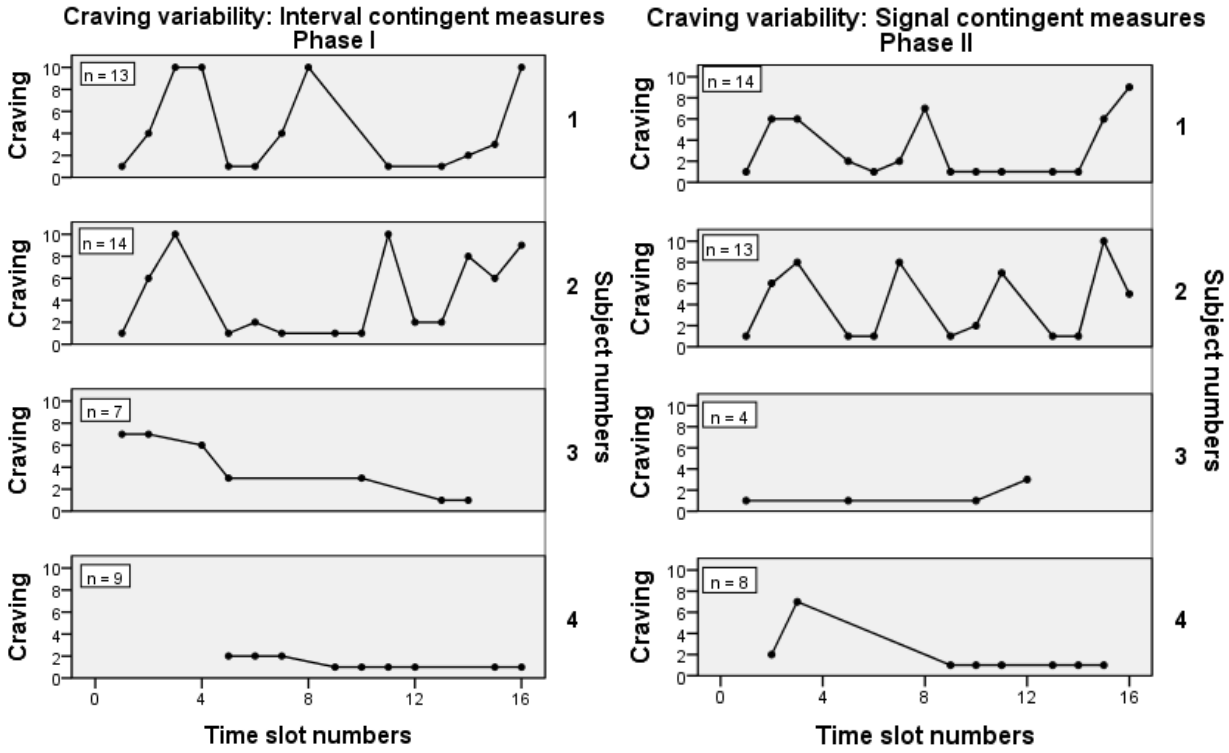

Figure 2. Visualisation of craving scores per participant over time. The missing values are left out, a line has been drawn through the recordings. The possible craving scores range from 1-10. The time slot numbers are displayed on the horizontal axe. Every four numbers represent one day: First time slot = morning, second time slot = afternoon, third and fourth time slots are evening.

4.2 Perceived burden and compliance of the IRM-NS designs

4.2.1 Perceived burden of the IRM-NS designs

Another goal of the study was to explore the perceived burden of completing a questionnaire about alcohol craving multiple times a day. The questionnaire about the perceived burden the participants completed can be found in appendix D. The statements about the perceived burden of the participants were coded by the researcher (and coding by a second coder revealed a Cohen's kappa of 0,71 which corresponds with a substantial strength of agreement). The effort to complete the

questionnaire itself was generally experienced as being low. All participants thought the questionnaire itself was fairly 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"* (Subject 1, q.4).

The burden of filling in the questionnaires with a frequency of four times a day was perceived as relatively high by two participants.

> *"I thought it was annoying to receive text messages every time when there was a new questionnaire available.. I felt pressured to be available all the time"* (Subject 4, q.3).

Two participants reported feeling bad or stressed because of the questionnaires.

> *"I like serenity and to plan my time in my own manner. I don't like the feeling of always have 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"* (Subject 4, q.6).

> *"You have to keep in mind to answer on time, this caused stress"* (Subject 2, q.3)

However, it was also mentioned that it did not felt as a burden to answer the questions multiple times a day. The following answer was given to the question 'How much of a burden did you think it was to answer the questions multiple times a day?',

> *"Not really a burden, you have your phone often near you anyway."* (Subject 1, q.6).

The time of the day when the participants were prompted for answering the questions was sometimes mentioned. In general the alert at midnight was considered as unpractical because some participants were sleeping at that time and others did not have their mobile phone with them because they were with friends for example. The questionnaire at this time was sometimes perceived as being an extra 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).

One participant made a remark about the random and non-random time slots and preferred the time contingent measures.

*"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" (*Subject 1, q.2).

4.2.2 Compliance of the IRM-NS designs

Five occasions in total took place where equal to or more than 4 time slots in a row were missed, two in the time-contingent period and three in the signal-contingent period. The compliance for the time-contingent recordings was on average 67.19% (SD_{IC}=20.64) and for signal-contingent recordings 60,94% (SD_{SC}=29.01). Noticeable is that half of the participants completed a fairly high percentage of questionnaires (84,38% (SD=4.42) and 84,38% (SD=4.42) respectively of the time-contingent and signal-contingent recordings) and the other half fairly low: respectively 50,00% (SD=8.84) and 37,50% (SD=17.68).

When answering to the question 'Did you miss many timeslots? If yes, why?' the participants reported that sometimes they were too busy to react to the alert, were already asleep or had no internet connection.

"I missed a lot of time slots because I had no internet connection or was busy at my work" (Subject 4, q.11.).

4.3 Usability of the Empatica E4 wristlet

A sub goal of the study was to measure the usability of the E4 wristlet when wearing it on a daily basis for a week with use of the System Usability Scale. The participants filled in the System Usability Scale after completing the experiment. The average score was 65 points on a 100-point scale, this is considered as being below average. In Table 3 the average (converted) score the participants gave per item is visible, the score per participant can be found in appendix E. Item 1 (I think that I would like to use the E4 wristlet frequently) and 5 (I found the various functions in this E4 wristlet were well integrated) have a score lower than 2, this can be seen as low since the (converted) scores range from 1 to 4.

Table 3

Average given score converted per SUS item

System Usability Scale, Brooke (1996)	μ score
1. I think that I would like to use the E4 wristlet frequently	1.25
2. I found the E4 wristlet unnecessarily complex	2.75
3. I thought the E4 wristlet was easy to use	3.25
4. I think that I would need the support of a technical person to be able to use the E4 wristlet	4.00
5. I found the various functions in this E4 wristlet were well integrated	1.50
6. I thought there was too much inconsistency in the E4 wristlet	2.00
7. I would imagine that most people would learn to use the E4 wristlet very quickly	3.50
8. I found the E4 wristlet very cumbersome to use	2.00
9. I felt very confident using the E4 wristlet	2.25
10. I needed to learn a lot of things before I could get going with the E4 wristlet	3.50

Note. The converted SUS score ranges from 1-4

5. Discussion

The central aim of this study was to explore the feasibility of potential measurement methods for a PhD monitoring study that will monitor ex-alcohol addicts in their daily life and register their self reported psychological craving thoughts combined with measuring their physiological responses. The central aim was divided in three sub-goals.

5.1 Exploring the variability of craving measurements

The first sub-goal of this study was to explore the variability of measurements of alcohol craving with interval- and signal-contingent designs. The standard deviation of the average craving score measured with current study approaches the standard deviation of the average craving score Litt et al., (1998) measured in the laboratory with cue-exposure sessions, the distribution of craving and therefore the variability is very similar amongst both studies. This suggests that the designs used in the current study are able to measure craving in a comparable manner as in the laboratory.

The average craving that was reported in the current study for both designs together was considerable higher compared to EMA recordings of Litt et al., (1998) who reported a much lower mean craving score on a comparable scale ranging from 0-10 (instead of ranging from 1-10 as with this study). Also the frequency that some craving (>1) was reported in the current study is compared to other studies remarkably high. More than half of the recordings show craving in the current study. Litt et al., (1998) recorded only in 8% of the recordings a craving higher than zero. A highly likely reason for both of these differences is the different kind of participants that were used for the studies. Litt et al., (1998) measured craving in alcohol-dependent men who received treatment for their dependence, this study measured craving in heavily drinking not (yet) alcohol dependent students who had no intention to reduce their alcohol consumption. Whereas it is to be expected that the treated group of Litt et al., (1998) tried to avoid drinking cues, the participants of the current study did not. Also there is to be expected that individuals who sought treatment for their alcohol dependence are actively preventing the occurrence of alcohol craving, whereas the participants of this study did not. Lastly it is possible that alcohol dependent subjects are more reluctant to report their craving.

For both interval-contingent and signal-contingent recordings, two participants showed a high variability of craving and two participants had longer periods that their craving level was stable. One of the participants that had recordings which showed a less fluctuant craving reported that the craving she felt did not fluctuate very fast, this is in accordance with the measurements and suggests that the recordings have at least a reasonable validity.

The participants with a high average score for craving, scored high on the DSM-5 for an alcohol use disorder. This result is in accordance with Grusser, Morsen and Flor (2006), who state that the level of craving is higher in problem drinkers than it is with occasional drinkers. That these two design are able to (roughly) distinguish the level of craving between occasional and problem drinkers means the IRM-NS recordings show a similar craving pattern as could be expected.

Noticeable is that the average craving is relatively high in the last two time frames (evening) compared to the first two (morning and afternoon). Since craving is related in some research to higher levels of drinking (Ramirez & Miranda, 2014; Ray et al, 2010) this increased average of craving in the evening might be due to the highly likely possibility that students often drink alcohol in the evening and not during the

day (Kuntsche & Labhart, 2012). Treated alcohol addicts often will pursue abstinence or a controlled consumption of alcohol (Sobell, 2013), therefore a different pattern probably will appear in the craving measurements of ex-alcohol addicts.

The current study indicates that there is variability at all in craving levels measured by IRM-NS recordings. A 'craving pattern' is visible per person and whether this is similar to reality or not: it measures an aspect of craving in daily life. There can be concluded that interval-contingent as well as signal-contingent recordings are well suited for further use for measuring the variability of craving.

5.2 Perceived burden and compliance of the IRM-NS designs

5.2.1 Perceived burden of the IRM-NS designs

The second sub goal of this study was to explore the perceived burden of completing a questionnaire about alcohol craving multiple times a day for a week when using interval- and signal-contingent recordings. Little research could be found about the burden of the design as it was perceived by the participants themselves and therefore almost no research could be found on the influence of this perceived burden on the compliance. Most studies focus on the compliance rate and assume that if the compliance was high the burden was doable. In this study however, one participant did fill in the questionnaires but felt pressured. Even though the participant felt pressured she kept completing the tasks, otherwise she felt like she was not doing a good job. This implies that a high compliance is not necessary in accordance with a reasonable perceived burden.

The high frequency of completing the task was perceived as a burden by some participants. Some felt pressured because they had to be on time with answering the questions. Other participants did not perceive the burden as high, they thought it was 'slightly annoying' to fill in the questionnaires but did not consider it to be a burden. The burden of filling in the questionnaire at 12 midnight was too high, therefore it is not recommendable to require participants to fill in a time slot at midnight in future research. One participant made a difference between the interval and signal-contingent design in the burden she perceived. She preferred set times (interval-contingent) because she could anticipate for the questionnaires then. The effort of filling in the questionnaires itself and the difficulty of the questions themselves was seen as very low by all participants. Generally it can be concluded that none of the participants specifically enjoyed filling in the questionnaires but some did not mind

doing it whereas others perceived it as a burden. The burden they perceived ranged from neutral to negative. It is suggested to have a test phase first before the start of the actual data gathering to filter out the participants who perceive the burden as too high.

5.2.2 Compliance of the IRM-NS designs

The third sub goal of this study was to explore and compare the compliance of completing a questionnaire about alcohol craving multiple times a day for a week when using interval- and signal-contingent recordings. On average, roughly 60% (range 24-84%) of the tasks were completed in the week of the experiment for both recording types. Compared to similar research an average of 60% for completing the task is fairly low. In a study of Litt et al., (1998), on average 81% (range 49-100%) of the time participants completed the tasks with a signal-contingent design. In another comparable study, 79% of the random prompt assessments were completed (Tidey et al., 2008). The low average can be explained partially due to the fact that both the participants with the lowest compliance rates only had access to internet via WiFi. This means that these participants were not always able to fill in the questionnaire because sometimes there was no WiFi. For future research the compliance for interval- and signal-contingent recordings can be increased by making sure every participant has access to internet all day. Furthermore, a possible explanation for the high compliance rate in the study of Litt et al., (1998) is the fact that the participants had an external motivation for completing the tasks: participants received 35 dollars at the end of the week of the experiment. Participant motivation plays next to the perceived burden of the study an important role in the degree of compliance (Green, Rafaeli, Bolger, Shrout & Reis., 2006). To increase the compliance it is important to create and maintain participant motivation (Christensen, Barrett, Bliss-Moreau, Lebo, Kaschub., 2003). Participant motivation can be increased by rewarding the participants if their compliance rate is high, such as Beckham et al. (2008) suggested in a previous study. A higher compliance rate can also be accomplished by the research team having a positive attitude and the establishment of a good working relationship between researchers and participants (Christensen et al., 2003).

A factor that plays an important role in the degree of compliance in studies with a high participant demand is the burden created by the study (Green, et al., 2006). As mentioned before, a high burden can result in high attrition rates and noncompliance (Moskowitz et al., 2009). The participants that reported a relatively

high perceived burden have a compliance lower than that of the participants that reported the perceived burden as doable. Furthermore, it is plausible that the two participants who reported a higher perceived burden had a more negative attitude towards filling in the daily questionnaires than the other two participants. It is likely that this negative attitude decreased the compliance rate because participants who feel negative about the study are likely to be less compliant. To create and maintain a positive attitude with the participants towards filling in the daily questionnaires, they should have the feeling that their contribution is important and that they are doing something meaningful with their time (Conner & Bliss-Moreau, 2005). An important foot note is that in the current study the participants who had a low compliance only had access to internet via WiFi, this is the probably the primary reason of these two participants having a relative low compliance in this particular study.

5.3 Usability of the Empatica E4 wristlet

The fourth sub-goal of this study was to measure the usability of the Empatica E4 wristlet with the System Usability Scale when wearing it on a daily basis for a week. With the System Usability Scale the usability of the E4 wristlet is rated slightly below average by the participants. According to the adjective rating scale described by Bangor, Kortum and Miller, (2009) this means that the E4 wristlet has a 'good' usability, but it scores in the lower spectrum of 'good'. Item 1 (I think that I would like to use the E4 wristlet frequently) and 5 (I found the various functions in this E4 wristlet were well integrated) had a relatively low score. The low score on item 1 can possibly be explained due to the comments the participants gave on the comfort of the wristlet. Some participants did not think the wristlet was comfortable at all, which could explain why they gave a low score when asked if they would like to use the E4 wristlet frequently. A possible reason for item 5 (I found the various functions in this E4 wristlet were well integrated) to be low is the fact that the E4 wristlet probably did not have various functions from the perspective of the participants. The participants only had to turn the wristlet on so that it worked and to upload the data. It is possible that the participants gave a low score to this item because there are basically no various functions to be integrated. Considering only the usability, the E4 wristlet seems to score good enough on the SUS to be used in the monitoring study. However, since some participants reported a low level of comfort of the bracelet, it is recommended to explore the wearability of the wristlet.

5.4 Limitations of the study

The design and execution of this study has several limitations. A possible limitation of this study is the small sample size: with four participants no conclusions on population levels can be made. However, Molenaar and Campbell (2009) state that inferences about the state of affairs at populations levels rarely can be equated at intraindividual (within person) level. They state that analysis of intraindividual variation should be done to optimally guide the psychological processes concerned. This time series models, where multiple measures are done within a person, can be used to carry out feedback-feed forward guidance in real time (Molenaar & Campbell, 2009). This intraindividual research provides a good start for the eventual feedback system which is aimed at for the PhD study this feasibility study is done for.

Another limitation is that the order in which the designs were offered was not varied, all the participants first experienced the interval contingent design and after that the signal contingent design. Because there was not varied in the order it is possible that a bias occurred such as improvement in performance due to repeated practise, habituation to the task or a decline in performance because of the many repeated questionnaires (Cozby, 2009). To avoid a bias due to order effects an option is to alternate the order of the two designs per participant to counterbalance the possible bias due to the order in which the designs are offered.

In this study only self reported craving was taken into consideration. The reliability of self reported craving is sometimes questioned, because some conditions where strong urges are expected often do not reveal high levels of craving when self reported (Wertz & Sayette, 2001). Furthermore, some researchers suggest that the 'wanting system', which underlies the concept of craving, operates primarily at implicit, unconscious levels and that this only sometimes comes into the conscious awareness of a person (Robinson & Berridge, as mentioned by Tiffany & Wray, 2014). Because of this possible unconscious craving, the question arises if the measurement of self reported craving captures all craving that a person experiences or only a part of it. However, even if self reported craving is only the top of the iceberg, it is not evident how one can capture unconscious craving without relying on a explicit self-report measures about the user's experience of craving (Tiffany & Wray, 2014).

The participant group consisted of four students who did not have an internal or external motivation to fill in the questionnaires. Because of this it is possible that

the participants experienced the perceived burden of filling in the questionnaires different than individuals with internal motivation would. Participants having a sufficient motivation is essential to all studies involving self reports (Green et al., 2006). Furthermore, because the participant group did not consist out of ex alcoholics, a possibly important factor that is not taken into consideration is the occurrence of withdrawal symptoms. These symptoms that can occur after stopping with drinking alcohol could influence the reported craving and perceived burden, for example because of the significant stress or impairment in social, occupational or other important areas as described in the DMS-5. It is possible that participants do not complete recordings on a particular stressful moment for example (Larson & Almeida, 1999).

At last, a factor that might have influenced the perceived burden of the participants is the fact that the researcher herself had to text the alert to the participants. This was because the used app for the questionnaire did not work properly. A personal text is probably seen as more pressing than a non personal alert from an app. This could have caused participants to report a higher perceived burden because they felt more pressured than would have been the case with a non personal alert.

5.5 Recommendations for further research and conclusion

The current study has positive results regarding the use of IRM-NS recordings for measuring the variability of craving. It seems that the variability of the craving measurements does not differ much between the designs, however it is recommended for the PhD monitoring study to choose for an interval-contingent design. With an interval-contingent design participants can anticipate better on the questionnaires and therefore they can integrate the daily questionnaires better in their life. Also the burden of an interval-contingent design is generally lower (Conner & Bliss-Moreau, 2005). It is recommended to have a test phase with participants before the start of the actual study. This makes it possible for the participants who experience the design too much as a burden to drop out with researchers still having time to find replacing participants. Furthermore, the Empatica E4 wristlet seems suitable to be used during the PhD monitoring study. However, the possible low wearability is an issue that might cause some participants to drop out.

This research is an example of a study on intraindividual level for the exploring of measuring self-reported craving in daily life and can give direction to further research for as well the measurement of craving as the use of intensive repeated measures in a naturalistic setting in general. Recommended is to do research on an intraindividual level more often in order to clarify the pattern of craving within the individual. Because after all, in the end it is the individual that matters.

6. References

Alliger, G. M., & Williams, K. J. (1993). Using signal-contingent experience sampling methodology to study work in the field: A discussion and illustration examining task perceptions and mood. *Personnel Psychology, 46*(3), 525-549.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction*, *24*(6), 574-594.

Beckham, J. C., Wiley, M. T., Miller, S. C., Dennis, M. F., Wilson, S. M., McClernon, F. J., & Calhoun, P. S. (2008). Ad lib smoking in post-traumatic stress disorder: An electronic diary study. *Nicotine & Tobacco Research*, *10*(7), 1149-1157.

Brooke, J., Bevan, N., Brigham, F., Harker, S., & Youmans, D. (1990, August). Usability statements and standardisation: Work in progress in ISO. In *Proceedings of the IFIP TC13 Third International Conference on Human-Computer Interaction* (pp. 357-361). North-Holland Publishing Co.

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual review of psychology*, *54*(1), 579-616.

Bottlender, M., & Soyka, M. (2004). Impact of craving on alcohol relapse during, and 12 months following, outpatient treatment. *Alcohol and Alcoholism*,*39*(4), 357-361.

Christensen, T. C., Barrett, L. F., Bliss-Moreau, E., Lebo, K., & Kaschub, C. (2003). A practical guide to experience-sampling procedures. *Journal of Happiness Studies*, *4*(1), 53-78.

Conner, T., Bliss-Moreau, E. (2005). Sampling Human Experience in Naturalistic Settings. In Hesse-Biber, S. N., & Leavy, P. (Eds.), *Emergent Methods in Social Research* (pp. 109-131). Thousand Oaks, CA: SAGE Publications

Cozby, P. C. (2009). Methods of Behavioral Research: Tenth Edition. New York, NY: McGraw-Hill.

Dawson, D. A., Goldstein, R. B., & Grant, B. F. (2007). Rates and Correlates of Relapse Among Individuals in Remission From DSM-IV Alcohol

Dependence: A 3-Year Follow-Up. *Alcoholism: Clinical and Experimental Research*, *31*(12), 2036-2045.

Drobes, D. J., & Thomas, S. E. (1999). Assessing craving for alcohol. *Alcohol Research & Health*, *23*(3), 179-179.

Evren, C., Durkaya, M., Evren, B., Dalbudak, E., & Cetin, R. (2012). Relationship of relapse with impulsivity, novelty seeking and craving in male alcohol-dependent inpatients. *Drug and alcohol review*, *31*(1), 81-90.

Flannery, B. A., Roberts, A. J., Cooney, N., Swift, R. M., Anton, R. F., & Rohsenow, D. J. (2001). The role of craving in alcohol use, dependence, and treatment. *Alcoholism: Clinical and Experimental Research*, *25*(2), 299-308.

French, D. P., & Sutton, S. (2010). Reactivity of measurement in health psychology: How much of a problem is it? What can be done about it? *British journal of health psychology*, *15*(3), 453-468.

Gordon, S. M., Sterling, R., Siatkowski, C., Raively, K., Weinstein, S., & Hill, P. C. (2006). Inpatient desire to drink as a predictor of relapse to alcohol use following treatment. *American Journal on Addictions*, *15*(3), 242-245.

Green, A. S., Rafaeli, E., Bolger, N., Shrout, P. E., & Reis, H. T. (2006). Paper or plastic? Data equivalence in paper and electronic diaries. *Psychological methods*, *11*(1), 87.

Grüsser, S. M., Mörsen, C. P., & Flor, H. (2006). Alcohol craving in problem and occasional alcohol drinkers. *Alcohol and Alcoholism*, *41*(4), 421-425.

Hufford, M. R., & Shiffman, S. (2002). Methodological issues affecting the value of patient-reported outcomes data. *Expert review of pharmacoeconomics & outcomes research*, *2*(2), 119-128.

Jellinek. (n.d.). Alcoholverslaving. Retrieved from https://www.jellinek.nl/ben-ik-verslaafd/alcoholverslaving/

Kuntsche, E., & Labhart, F. (2012). Investigating the drinking patterns of young people over the course of the evening at weekends. *Drug and alcohol dependence*, *124*(3), 319-324.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Larson, R. W., & Almeida, D. M. (1999). Emotional transmission in the daily lives of families: A new paradigm for studying family process. *Journal of Marriage and the Family*, 5-20.

Litt, M. D., Cooney, N. L., & Morse, P. (1998). Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychology*, *17*(1), 48.

Molenaar, P. C., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112-117.

Moskowitz, D. S., Russell, J. J., Sadikaj, G., & Sutton, R. (2009). Measuring people intensively. *Canadian Psychology/Psychologie Canadienne*, *50*(3), 131.

Ramirez, J., & Miranda Jr, R. (2014). Alcohol craving in adolescents: bridging the laboratory and natural environment. *Psychopharmacology*, *231*(8), 1841-1851.

Ray, L. A., Miranda Jr, R., Tidey, J. W., McGeary, J. E., MacKillop, J., Gwaltney, C. J., & Monti, P. M. (2010). Polymorphisms of the μ-opioid receptor and dopamine D_4 receptor genes and subjective responses to alcohol in the natural environment. *Journal of abnormal psychology*, *119*(1), 115.

Rehm, J. Ü. R. G. E. N., Shield, K. D., Rehm, M. X., Gmel, G., & Frick, U. (2012). Alcohol consumption, alcohol dependence and attributable burden of disease in Europe. *Centre for Addiction and Mental Health*.

Schwarz, N. (2012). Why Researchers Should Think "Real-Time" - A Cognitive Rationale. In Mehl M. R., & Conner T. S. (Eds.), *Handbook of research methods for studying daily life* (pp. 22-39). New York: Guilford Press.

Scollon, C. N., Prieto, C. K., & Diener, E. (2009). Experience sampling: promises and pitfalls, strength and weaknesses. In *Assessing well-being* (pp. 157-180). Springer Netherlands.

Smyth, J., Wonderlich, S., Crosby, R., Miltenberger, R., Mitchell, J., & Rorty, M. (2001). The use of ecological momentary assessment approaches in eating disorder research. *International Journal of Eating Disorders*, *30*(1), 83-95.

Sobell, M. (2013). *Behavioral treatment of alcohol problems: Individualized therapy and controlled drinking* (Vol. 563). Springer Science & Business Media.

Takarangi, M., Garry, M., & Loftus, E. (2006). Dear diary, is plastic better than paper? I can't remember: Comment on Green, Rafaeli, Bolger, Shrout, and Reis (2006). *Psychological Methods, 11*, 199–122.

Tidey, J. W., Monti, P. M., Rohsenow, D. J., Gwaltney, C. J., Miranda, R., McGeary, J. E., & Paty, J. A. (2008). Moderators of Naltrexone's Effects on Drinking, Urge, and Alcohol Effects in Non-Treatment-Seeking Heavy Drinkers in the

Natural Environment. *Alcoholism: Clinical and Experimental Research*,*32*(1), 58-66.

Tiffany, S. T. (1992). A critique of contemporary urge and craving research: Methodological, psychometric, and theoretical issues. *Advances in Behaviour Research and Therapy*, *14*(3), 123-139.

UNDCP/WHO Informal Expert Committee on the Craving Mechanism (1992), United Nations International Drug Control Programme and World Health Organization Technical report series No. V.

Wertz, J. M., & Sayette, M. A. (2001). A review of the effects of perceived drug use opportunity on self-reported urge. *Experimental and clinical psychopharmacology*, *9*(1), 3.

Wheeler, L., & Reis, H. T. (1991). Self-recording of everyday life events: Origins, types, and uses. *Journal of personality*, *59*(3), 339-354.

Wray, T. B., Merrill, J. E., & Monti, P. M. (2014). Using ecological momentary assessment (EMA) to assess situation-level predictors of alcohol use and alcohol-related consequences. *Alcohol research: current reviews*, *36*(1), 19.

7. Appendix

7.1 Appendix A - Randomized timeslots signal contingent design

Timeslot	Randomized time	Timeslot	Randomized time
1	10:45 am	9	11:30 am
2	02:00 pm	10	03:45 pm
3	05:30 pm	11	06:00 pm
4	12 midnight	12	12 midnight
5	09:00 am	13	09:30 am
6	01:30 pm	14	03:15 pm
7	05:00 pm	15	08:45 pm
8	12 midnight	16	12 midnight

7.2 Appendix B - Questionnaire questions

- Hoeveel glazen alcohol heb je gister gedronken? (Amount of alcohol)
- Hoeveel trek heb je in alcohol op dit moment? (Craving)
- Hoeveel vertrouwen heb je op dit moment dat je alcohol kan weerstaan, zelfs als je trek voelt? (Trust to refuse alcohol)
- Hoe voel je je op dit moment? (Feeling rated with availance-arousal scale)

7.3 Appendix C - Informed consent

Titel onderzoek: Alcohol in het dagelijks leven

Verantwoordelijke onderzoeker: Jessica Stroes

In te vullen door de deelnemer

Ik verklaar op een voor mij duidelijke wijze te zijn ingelicht over de aard, methode, doel. Ik weet dat de gegevens en resultaten van het onderzoek alleen anoniem en vertrouwelijk aan derden bekend gemaakt zullen worden. Mijn vragen zijn naar tevredenheid beantwoord.

Ik stem geheel vrijwillig in met deelname aan dit onderzoek. Ik behoud me daarbij het recht voor om op elk moment zonder opgaaf van redenen mijn deelname aan dit onderzoek te beëindigen.

Naam onderzoeker.....

Datum.....Handtekening onderzoeker.....

In te vullen door de uitvoerende onderzoeker

Ik heb een mondelinge en schriftelijke toelichting gegeven op het onderzoek. Ik zal resterende vragen over het onderzoek naar vermogen beantwoorden. De deelnemer zal van een eventuele voortijdige beëindiging van deelname aan dit onderzoek geen nadelige gevolgen ondervinden.

Naam onderzoeker.....

Datum.....Handtekening onderzoeker.....

7.4 Appendix D - Questionnaire perceived burden

Participant 1

1. Wat vond je ervan om de vragenlijst iedere dag in te vullen? Geef uitleg
Geen probleem, de vragenlijst had je zo ingevuld.

2. Zou je het (dagelijks) invullen van de vragenlijst kunnen integreren in je dagelijks leven? Zo ja, waarom wel of niet?
Ja, alhoewel het wel gemakkelijker is als het vaste tijden zijn, dan kan je daarop anticiperen en vergeet je het naar verloop van tijd ook niet meer.

3. Wat is je houding ten opzichte van de dagelijkse surveys? Geef uitleg
Mijn houding tegenover de dagelijkse surveys is vrij positief; het kostte weinig tijd om in te vullen, waren gemakkelijke vragen waar je niet te lang over na hoefde te denken en via de app was het zo gebeurd. Alleen als de vragen later op de avond/ s nachts waren is het lastig ze in te vullen, ook als je nog wakker bent (dan denk je er niet meer aan op dat tijdstip, ik in ieder geval niet zo snel).

4. Hoe ging het steeds beantwoorden van de vragen? Geef uitleg
Het beantwoorden van de vragen ging goed, het waren makkelijke vragen waar vrij genuanceerd, maar toch makkelijk, in te vullen is hoe jij erover denkt op dat moment.

5. Hoe ingewikkeld vond het het beantwoorden van de vragen? Geef uitleg
Het beantwoorden van de vragen was totaal niet ingewikkeld, zelfs vrij gemakkelijk.

6. Hoe belastend vond je het om meerdere keren per dag de vragen te beantwoorden? Geef uitleg
Niet zo belastend, je telefoon heb je vaak toch bij de hand. Het is wel fijn om een herinnering gestuurd te krijgen, anders vergeet je het snel.

7. Irriteerde je je ergens aan bij het beantwoorden van de vragen? Zo ja, waarom?
Ja, vooral bij de laatste vraag (over energie/humeur) hangt best wel los van de andere vragen, wat wel een vertekend beeld kan geven (maar misschien lag dit aan mij).

8. Zijn er dingen mis gegaan bij het beantwoorden van de vragen? Zo ja, wat en waarom?
Een keer deed de app het niet en ik ben het vast een keer vergeten in te vullen (omdat ik met wat anders bezig was of omdat ik sliep).

9. Hoe vond je het om de vragen te beantwoorden op de app? Geef uitleg
Ik vond het prima; het kostte niet veel tijd en waren gemakkelijke vragen.

10. Vond je de app makkelijk te gebruiken? Zo ja, waarom wel/niet?
Ja, het werkte snel en je hebt je telefoon vaal bij de hand, dus dat is makkelijk.

11. Heb je veel invulmomenten voor de vragenlijsten gemist? Zo ja, waarom?

Een paar denk ik, omdat ik of lag te slapen of er 's avonds laat niet aan dacht (heb dan meestal geen telefoon bij de hand). Misschien overdag ook wel een keertje omdat ik bijvoorbeeld druk aan het studeren was.

Participant 2

1. Wat vond je ervan om de vragenlijst iedere dag in te vullen? Geef uitleg
Dat was prima te doen, soms was ik wel zo druk aan het werk dat ik het vergat (gebruik web what's app dus dan kijk je niet zo vaak op je telefoon) en de vragenlijst om 12u 's nachts was soms wat onhandig (ik was of op de kroeg zonder telefoon of lag al te slapen). Het duurde niet lang dus was niet storend. Misschien zou het handig zijn als op het horloge een lichtje gaat branden oid dat je het moet invullen zodat je er duidelijk aan herinnert wordt.

2. Zou je het (dagelijks) invullen van de vragenlijst kunnen integreren in je dagelijks leven? Zo ja, waarom wel of niet?
Ik denk dat ik er wel aan zou kunnen wennen maar het is vooral lastig dat je het soms gewoon vergeet en / of geen telefoon bij je in de buurt hebt. Met werk en / of sporten gebruik ik mijn telefoon niet dus dat kan lastig zijn.

3. Wat is je houding ten opzichte van de dagelijkse surveys? Geef uitleg
Je moet er wel op letten dat je ze op tijd beantwoord, dat zorgt voor stress. Ook als ik om half 12 wilde gaan slapen zat ik zo van, oh, nu kan ik net niet die vragenlijst invullen, zou ik nu wakker moeten blijven? etc

4. Hoe ging het steeds beantwoorden van de vragen? Geef uitleg
Op een gegeven moment waren de vragen niet meer leesbaar in de app (hadden bijna dezelfde kleur als de achtergrond) en één keer was het niet mogelijk om de vraag in te vullen (stond er alleen 'infinity'). Verder merkte ik dat ik de vragen na een tijdje uit mijn hoofd wist.

5. Hoe ingewikkeld vond het het beantwoorden van de vragen? Geef uitleg
Het is best lastig om aan te geven hoeveel trek je hebt in alcohol. Daarnaast vond ik de grafiek met plezierig/onplezierig en hoog in energie/laag in energie lastig,

aangezien ik de afgelopen week heel erg moe was en het daarom lastig was in te vullen (anders staat er altijd 'laag in energie', dus ik moest voor mezelf een nieuwe schaal bedenken zegmaar).

6. Hoe belastend vond je het om meerdere keren per dag de vragen te beantwoorden? Geef uitleg
Niet erg belastend, alleen vervelend als je met mensen in gesprek bent of ergens mee bezig bent en ineens dat lijstje weer in moet vullen. Ik was vaak bang dat ik te laat zou zijn zodat de lijst niet meer beschikbaar zou zijn.

7. Irriteerde je je ergens aan bij het beantwoorden van de vragen? Zo ja, waarom?
Dat het lastig was om een antwoord te geven op de plezierig/onplezierig en hoog in energie/laag in energie

8. Zijn er dingen mis gegaan bij het beantwoorden van de vragen? Zo ja, wat en waarom?
Wanneer je bij de grafiek in wilde vullen hoe je je voelde en vervolgens op het knopje wilde drukken om de lijst af te ronden (het groene vinkje), dan ging het andere groene vinkje van hoe je je voelde vaak ineens mee naar die hoek en moest ik het weer opnieuw invullen. Sowieso reageerde het vinkje niet snel, je moest echt je vinger langere tijd op een plek in de grafiek drukken voordat ie verscheen. Ook waren de vragen op een gegeven moment niet meer goed leesbaar.

9. Hoe vond je het om de vragen te beantwoorden op de app? Geef uitleg
Prima, is al wel beschreven in de andere vragen.

10. Vond je de app makkelijk te gebruiken? Zo ja, waarom wel/niet?
Ja, hij gaf zelf meldingen als ik m in moest vullen. Alleen dat vinkje wat niet werkte (zie vraag 8) was vervelend. Daarnaast was ie een keer uit zichzelf uitgelogd, dat was onhandig.

11. Heb je veel invulmomenten voor de vragenlijsten gemist? Zo ja, waarom?
Ja vooral die van 12u 's avonds, doordat ik dan al sliep of op de kroeg was (waar ik nooit een telefoon bij me heb)

Participant 3

1. Wat vond je ervan om de vragenlijst iedere dag in te vullen? Geef uitleg
Vond het niet leuk, maar ook niet moeilijk. Is gewoon jammer omdat ik niet altijd internet om mee heen had zodat ik der vragenlijst alleen soms kon invullen.. En ook dat lukte met internet niet zo goed.. Vond de laatste vraag altijd misgaan

2. Zou je het (dagelijks) invullen van de vragenlijst kunnen integreren in je dagelijks leven? Zo ja, waarom wel of niet?

Nee want ik vind dat heel lastig zie de punten bij 1). Ik denk ook dat je daar heel snel kan liegen .. Verder als je een vaste baan hebt kun je niet altijd op je mobiel of wat dan ook letten

3. Wat is je houding ten opzichte van de dagelijkse surveys? Geef uitleg
Zou ik heel erg vinden omdat het dan echt mijn dagelijks leven zou beïnvloeden en ik op vaste tijden zo iets moet doen.

4. Hoe ging het steeds beantwoorden van de vragen? Geef uitleg
Makkelijk. Dat was helemaal geen probleem en snel

5. Hoe ingewikkeld vond het het beantwoorden van de vragen? Geef uitleg
Helemaal niet want het waren heel makkelijke vragen en de wijze van invoeren was ook makkelijk

6. Hoe belastend vond je het om meerdere keren per dag de vragen te beantwoorden? Geef uitleg
Omdat ik er maar enkele moest doen wegens het gebrek aan internet vond ik het echt niet lastig of belastend maar als het zo was was het wel meestal zo dat ik een gesprek of zo even moest laten stil staan wat helemaal niet leuk was

7. Irriteerde je je ergens aan bij het beantwoorden van de vragen? Zo ja, waarom?
De laatste vraag met het grafiek. Kon mijn antwoord nooit terug zien

8. Zijn er dingen mis gegaan bij het beantwoorden van de vragen? Zo ja, wat en waarom?
Zie antwoord op vorig vraag .. Kan omdat het niet getoond werd

9. Hoe vond je het om de vragen te beantwoorden op de app? Geef uitleg
Best chill maar ook misschien te makkelijk omdat je dat dan ook met weinig aandacht kon doen

10. Vond je de app makkelijk te gebruiken? Zo ja, waarom wel/niet?
Ja wel.. Gewoon normale wijze van invoeren en makkelijk te begrijpen

11. Heb je veel invulmomenten voor de vragenlijsten gemist? Zo ja, waarom?
Ja wegens gebrek van internet verbinding als ik niet thuis ben

Participant 4

1. Wat vond je ervan om de vragenlijst iedere dag in te vullen? Geef uitleg
Ik vond het best wel lastig om eerlijk te zijn. Ook als het maar 4 keer per dag was von ik het veel. Ook had ik niet overaal wifi zodat het niet mogelijk was om de vraagenlijst in te vullen.

2. Zou je het (dagelijks) invullen van de vragenlijst kunnen integreren in je dagelijks leven? Zo ja, waarom wel of niet?

Niet echt. Vaak ben je in situaties waar het stoort of niet kan. Bijvorbeeld op school of werk. Ook zo als al angegeven heb in niet overaal wifi om het in te vullen.

3. Wat is je houding ten opzichte van de dagelijkse surveys? Geef uitleg
Ik vond het irritant steeds smsmjes te krijgen dat er een nieuwe survey beschikbaar is. ik voelde me onder druk gezet dat ik steeds berijkbaar moest zijn.

4. Hoe ging het steeds beantwoorden van de vragen? Geef uitleg
Slecht omdat ik vaak niet eens de mogelijkheid had om het de doen vanwegen die wifi en tijd.

5. Hoe ingewikkeld vond het het beantwoorden van de vragen? Geef uitleg
Dat vond ik helemaal niet ingewikkelt. Het waren altijd de zelfde vraagen zodat ik het soms lastig vond omdat zich binnen paar uur niet veel heeft veranderd.

6. Hoe belastend vond je het om meerdere keren per dag de vragen te beantwoorden? Geef uitleg
Best wel lastig. Ik hou van rust en davon mijn tijd zelf in te plennen. Ik vind het gevoel altijd iets te moeten beantwoorden niet leuk. Ik voel me erg onder druk. Ook als ik vraagen niet kunt beantwoorden voelde ik me slecht omdat ik mijn oogaven niet heb gedaan.

7. Irriteerde je je ergens aan bij het beantwoorden van de vragen? Zo ja, waarom?
Ja dat het zo vaak was en altijd het zelfde. Ik kont mijn gevoelens ook niet goed inschatten . Dus soms heb ik maar wat aangegeven omdat ik het niet echt wisste.

8. Zijn er dingen mis gegaan bij het beantwoorden van de vragen? Zo ja, wat en waarom?
Ja omdat het altijd het zelfde was en mijn gevoelens niet zo snel veranderen. Ook waren de vraagen beetje vaag.

9. Hoe vond je het om de vragen te beantwoorden op de app? Geef uitleg
Lastig omdat het zo vaak was.

10. Vond je de app makkelijk te gebruiken? Zo ja, waarom wel/niet?
Ja het was heel makkelijk omdat het heel oversichtelijk was. Ook waren het niet veele of ingewikkelde vraagen.

11. Heb je veel invulmomenten voor de vragenlijsten gemist? Zo ja, waarom?
Ja vaaker omdat ik geen wifi had of heen tijd omdat ik op werk sat.

7.5 Appendix E - System Usability Scale scores

	Subject 1	Subject 2	Subject 3	Subject 4
1.I think that I would like to use the E4 wristlet frequently	2	2	1	4
2. I found the E4 wristlet unnecessarily complex	1	3	3	2
3. I thought the E4 wristlet was easy to use	4	5	3	5
4. I think that I would need the support of a technical person to be able to use the E4 wristlet	1	1	1	1
5. I found the various functions in this E4 wristlet were well integrated	3	1	3	3
6. I thought there was too much inconsistency in the E4 wristlet	3	3	2	4
7. I would imagine that most people would learn to use the E4 wristlet very quickly	5	4	4	5
8. I found the E4 wristlet very cumbersome (frustrerend) to use	4	4	3	1
9. I felt very confident using the E4 wristlet	3	2	4	4
10. I needed to learn a lot of things before I could get going with the E4 wristlet	1	3	1	1

7.6 Appendix F - Questionnaire alcohol consumption

	Subject 1	Subject 2	Subject 3	Subject 4
Geslacht	Vrouw	Vrouw	Vrouw	Vrouw
Leeftijd	20	22	23	24
Hoe veel glazen alcohol drink je gemiddeld per week?	18-22	>24	8-12	8-12
Hoe vaak drink je gemiddeld per week?	5-6	5-6	2-3	2-3
Hoe vaak per maand drink je gemiddeld?	>6	>6	4-5	>6
Hoe vaak per maand drink je meer dan 6 glazen alcohol?	>6	>6	3-4	2-3

7.7 Appendix G - DSM-5 Alcohol Use Disorder

	Subject 1	Subject 2	Subject 3	Subject 4
1. Had times when you ended up drinking more, or longer, than you intended	Yes	Yes	No	Yes
2. More than once wanted to cut down or stop drinking, or tried to, but couldn't?	No	Yes	No	Yes
3. Spent a lot of time drinking? Or being sick or getting over other aftereffects?	Yes	Yes	No	Yes
4. Wanted a drink so badly you couldn't think of anything else?	No	No	No	No
5. Found that drinking—or being sick from drinking—often interfered with taking care of your home or family? Or caused job troubles? Or school problems?	Yes	Yes	No	No
6. Continued to drink even though it was causing trouble with your family or friends?	No	No	No	No
7. Given up or cut back on activities that were important or interesting to you, or gave you pleasure, in order to drink?	No	No	No	No
8. More than once gotten into situations while or after drinking that increased your chances of getting hurt	No	No	No	Yes

(such as driving, swimming, using machinery, walking in a dangerous area, or having unsafe sex)?				
9. Continued to drink even though it was making you feel depressed or anxious or adding to another health problem? Or after having had a memory blackout?	Yes	Yes	No	Yes
10. Had to drink much more than you once did to get the effect you want? Or found that your usual number of drinks had much less effect than before	Yes	Yes	No	No
11. Found that when the effects of alcohol were wearing off, you had withdrawal symptoms, such as trouble sleeping, shakiness, restlessness, nausea, sweating, a racing heart, or a seizure? Or sensed things that were not there?	No	No	No	No
Total questions answered with yes	5	6	0	5

7.8 Appendix H - Participant Instruction

Procedure

Welkom bij het onderzoek 'Alcohol in het dagelijks leven'. Allereerst hartelijk bedankt voor je deelname! Er volgt een korte uitleg over het onderzoek en de taken die moeten worden uitgevoerd om het onderzoek te voltooien.

Dit is een onderzoek over het verlangen naar alcohol in het dagelijks leven. In dit onderzoek voer je diverse taken uit gedurende 8 dagen. Aan het eind van het onderzoek is er een interview over je ervarigen tijdens het onderzoek met de manier van meten en de gebruikte materialen. Jou bijdragde in dit onderzoek onderzoek bestaat uit 3 onderdelen

1. Een aantal keer per dag een vragenlijst invullen in combinatie met een korte blaastest (Hoofdonderzoeker is Jessica Stroes).
2. Het dragen van een armband die je fysiologische waarden constant meet (Hoofdonderzoeker is Niklas Enewoldsen).
3. Het beantwoorden van vragen over je ervaringen tijdens een interview achteraf (Hoofdonderzoeker is Mira Oberhagenmann).

Over ieder onderdeel zal hieronder meer uitleg volgen. Als je tijdens het onderzoek vragen hebt of iets niet begrijpt is de hoofdonderzoeker van dat onderdeel altijd

telefonisch bereikbaar. Ook zit je in een groepsapp met de onderzoekers en andere participanten, hier kun je altijd vragen stellen of opmerkingen maken. Het is belangrijk om te weten dat dit een pilot onderzoek is. Er kunnen zaken mislopen of onduidelijkheden optreden waar de onderzoekers geen rekening mee hebben gehouden. Als dit het geval is, is het van belang om contact op te nemen met een van de onderzoekers. Er kunnen dan eventueel zaken worden aangepast of uitgelegd.

Uitleg dagelijkse vragenlijst

In dit deel van het onderzoek onderzoek krijg je iedere dag een aantal vragen gesteld, in de ochtend, middag en twee keer in de avond. Aan het eind van iedere vragen 'sessie' is de bedoeling om een korte blaastest uit te voeren. Het zijn korte vragen en de gehele 'taak' zal niet langer duren dan een paar minuten.

De eerste vier dagen van het onderzoek zal je iedere dag een reminder bericht krijgen dat je de vragenlijst moet invullen. Dit zal iedere dag op de dezelfde tijden zijn: 11:00, 15:00, 19:00 en 24:00. De vijfde tot en met de achtste dag zal je ook vier keer per dag een reminder krijgen, dit keer krijg je echter drie van de vier keer op random tijden een reminder. Deze random tijden vallen wel binnen een bepaald tijdframe. De tijdframes zijn als volgt: 9:00 - 13:00, 13:00 - 17:00, 17:00 - 21:00. De vierde reminder krijg je, net als de eerste vier dagen, om 24:00. Omdat niet iedereen nog wakker is om 24:00 is het begrijpelijk dat het niet (altijd) lukt om dit tijdslot te halen. Het is dan ook niet erg als je het laatste moment van de dag mist. Op het moment dat je een alert krijgt voor een sessie is het belangrijk dat je de sessie uitvoert zodra dit kan. Je hebt twee uur de tijd om het uit te voeren, maar de voorkeur gaat er naar uit om de sessie zo snel mogelijk te voltooien na het krijgen van een alert.

De vragen zullen via een app worden gesteld. Hieronder staat een uitleg over het downloaden van de app en kun je oefenen met het uitvoeren van de vragen sessie.

UTSurvey app

1. Download de app 'UTSurvery' (van de maker Frank Borgonjen) op je smartphone. Deze app kun je downloaden bij 'Google Play' (Android) of bij de 'App Store' (Iphone).

2. Open de app op je smartphone en log in met de gegevens die je hebt gekregen van de onderzoeker.

3. Je ziet nu het begin scherm. Volg de instructies op het beginscherm, je krijgt nu een test vragenlijst te zien. Deze test vragenlijst is dezelfde vragenlijst die je iedere ochtend zult ontvangen.

4. Vul de test vragenlijst in. Het is belangrijk om een van de onderzoekers om opheldering te vragen als een vraag of iets anders niet helemaal duidelijk is.

Je bent nu klaar met het vragenlijst deel van de 'taak'.

Uitleg sensor

De E4 sensor is een draagbare biosensor die fysiologische reacties van de lichaam kan meten. De E4 komt met een USB-kabel en een adopter welk nodig is om de sensor op te laden. Om de data up te loaden is er een software nodig, namelijk de empatica manager. (ik had bedacht dat het misschien het handigst is om van te voren al een account aan te maken voor iedereen en dan tijdens de instructiedag hun deze gegevens te geven. Dit zou voor de proefpersoon makkelijker maken omdat zij of hij dan al de gegevens heeft en voor ons als onderzoekers omdat wij dan direct per empatica connect de upgeloade data kunnen inzien.) De sensor is niet watervast, dus voor het handen wassen, douchen enz. moet de sensor worden afgenomen. Tijdens het slapen hoeft de sensor niet gedragen te worden. De sensor moet elke avond worden opgeladen. Dit kan doormiddels hem in de stopcontact te stoppen of aan je laptop aan te sluiten.

Downloaden en installatie van de empatica manager

Om de data van de E4 up te loaden, is de empatica manager nodig (de volgende stappen moeten alleen één keer uitgevoerd worden):

1. installeer de empatica manager
2. sluit de empatica manager aan
3. koppel de E4 via USB aan je computer
4. verwijder de E4 uit de computer
5. start de empatica manager software en log in met jouw gegevens
6. koppel de E4 aan je computer

Uploaden van de data

De empatica manager wordt gebruikt om de data up te loaden. Darvoor moet je de software starten en je inloggen. De upload gebeurd dan automatisch en wordt bevestigd door een kleine berichtje. Deze procedure moet alle twee dagen gebeuren.

Gebruik van de E4

Om de sensor te starten moet je 1-2 seconden de button drukken. De E4 gaat automatisch in de record-mode welke zichtbaar wordt door een groene licht op de sensor. Bovendien is het mogelijk om markers te zetten door de button even kort te drukken (maar laat op dat je het niet te lang drukt en de sensor uitmaakt). Door de marker te zetten is het later mogelijk om de deze tijdpunt weer terug te vinden. Door de button 2 seconden te drukken, kan je de E4 uit schakelen.