UNIVERSITY OF TWENTE

An experience sampling study into intra-individual correlations between bodily signals and experienced feelings.

Jan-Luca Ahlemeyer

July 3rd 2019 - BMS Faculty - Department for Psychology

First Supervisor: Dr. M. Noordzij

Second Supervisor: Dr. M. Radstaak

Abstract

Investigations into the correlation between bodily states and perceived emotions may be beneficial in many medical sectors, however, to this point in time studies that investigated this relationship have mainly focused on inter-individual estimates. The aim of this study was to investigate the correlation between core affect and autonomic nervous system activity as measured by skin conductance levels. The skin conductance levels have been measured by wearable measurement devices (E4) during the daily activities of the test subjects, and the core affect state has been addressed with a mobile application (TIIM) that displays the two dimensions of core affect in one quadrant (core quadrant) in which the study participants were able to indicate their current state (i.e for the past minute) and the state of the last two hours. The data was analysed for both time intervals with a Spearman's rank-order correlation and due to the fact that emotional perceptions are highly influenced by individual factors, the analysis was done intra-individually as well as inter-individually and a comparison of the outcomes of both types of analysis has been made. The outcomes of both types of analysis intra-individually and inter-individually indicated that there likely is no linear correlation between skin conductance level and the dimensions of core affect. For the former a small amount of predominantly negative correlations has been detected between skin conductance level and the arousal dimension of core affect as well as between skin conductance level and the valence dimension; and for the latter, no significant correlation has been detected between skin conductance level and each of the two dimensions of core affect. These findings suggest that there likely is no direct linear relationship between bodily fluctuations and subjective experience in daily life, and all indicators for the presence of such a correlation indicate that it is negative (if existing at all) which is unexpected when comparing the results with other studies. Furthermore, by comparing the present study with similar research it can be assumed that individual and situational factors may play a crucial role in guiding the perceptions of emotions and autonomic nervous system activities may only play a small role.

Measurements of physiological signals as a means to estimate emotional states in humans can serve as a promising tool to assess or even to substitute self-report measures. Such developments may especially be promising in the medical sector when dealing with patients who are, due to for example verbal communication disabilities or other communication boundaries not capable of expressing their own emotional states (Picard, 2009; Noordzij, & Laroy-Noordzij, 2014). Moreover, it may also be useful to measure stress levels in tasks in which the individual is inhibited to articulate it him or herself due to the cognitive involvement in the task (Brouwer et al., 2018; Noordzij, Dorrestijn & Van den Berg, 2016). Finally also in case of a certain psychopathology such as borderline personality disorder (Derks, De Visser, Bohlmeijer, & Noordzij, 2017) or psychological personality traits such as Alexithymia, in which the individual is not able to monitor and express own emotions properly (Valdespino, Antezana, Ghane, & Richey, 2017), such somatic measurements may be efficient in supporting those affected to better understand their own affective states (de Bruin, Derks, & Noordzij, 2017). However, even though nowadays it is a widely accepted fact that subjective perceptions such as stress can have an impact on an individual's physical and mental health (Thoits, 2010) the nature of the correlation between subjective emotional perceptions and physiological signals, is yet to be explored regarding the within individual correlations in daily life, as will be explained later on. Therefore, the aim of the current study is to explore to which extent physiological signals are associated with emotional perceptions reflected by within individual correlations, and thus to investigate how useful body data is in predicting emotions. This investigation is determined by three characteristics: namely, the application of the theory of core affect by Russell (2009), a somatic measurement of emotions, and intra-individual estimates of emotions and their correlations with somatic measurements.

Issues with conceptions of emotion

The first issue one faces when intending to do research on emotion is the lack of a universal definition of the term “emotion” within the science of psychology. A lot of researchers have dealt with this problem during the last couple of centuries. Fehr and Russell concluded already in 1984 that no clear criteria can be set in order to define something as an emotion, but that states rather are labeled by humans as emotions depending on how much they feel the term “emotion” is applicable to the state in question (Fehr & Russell, 1984). Contemporary research highlights that the lack of a proper definition is the source of a lot of misconceptions, but also acknowledges that debates to this day have not been proven to be successful in creating a clear concept and that this is an aim not likely to be reached in future (Mulligan & Scherer, 2012).

Besides the unsuccessful efforts to establish a common understanding of the term emotion itself, also specific emotional states are not well defined. This is due to a number of reasons. For one, interpretation of emotions is to a large extent context dependent (Carroll & Russell, 1996; Fernández-Dols, Wallbott, & Sanchez, 1991), thus perceived states for example through bodily sensation can be the same but yet yield in different emotional perceptions for the individual. This makes the categorization of emotions very complicated. Moreover, within a given emotion, a high level of situational and individual variety can be observed (Russell, 2009) making it hard to construct an exclusively emotion-based framework that applies for different individuals. As can be concluded from these findings, traditional conceptions emotional states do not suffice as groundwork to start investigations into emotional research. In order to reduce the limitations of research on emotions, scientific investigations should apply a precise framework that defines emotional states, rather than using traditional emotion concepts.

Core affect

Core affect, according to Russell, is a neurological state that is the product of two dimensions, valence (also referred to as pleasure), and the arousal dimension. It is yet perceived as one unified impression by the individual. The valence dimension ranges from feeling pleasure or displeasure and the arousal dimension ranges from feeling energized to feeling torpid. The feeling that an individual gets as a result of these two dimensions is usually non-directive and is often caused by factors that lay outside of the individual's own consciousness (Russell, 2009). Because of that core affect is often the product of many influences and can be altered by real-life events, anticipated events, and drugs (Russell, 2009). Thus, field research appears to be at least equally as important as laboratory research and requires further investigation.

Russell does not view his theory on core affect as a concept that can be equated completely with the traditional view of emotions (2009); core affect can be rather viewed as a state that continues over time and is not characterized by a beginning or an end. In this sense, it shares similarities with nervous system activities. Therefore, ANS might be a useful form of somatic measurement for investigating a potential correlation between core affect and somatic activities. Because core affect can be viewed as transitioning process, and because recent studies investigated the relationship between core affect and somatic states often in the context of sudden stimuli exposure (Västfjäll & Gärling, 2007), long term investigations into this domain appear to be a much-needed contribution to this domain of research. This is also important as an application of devices that enhance the attention towards the own emotions (which are likely to be employed in such research) are theorized to increase awareness towards emotions over time (Farrell & Shaw, 1994).

Somatic aspects of research on emotions

While traditional emotion concepts could for the most part not be successfully linked to specific activity in the autonomic nervous system (ANS) (Feldman Barrett, 2008) and while changes in the ANS can be explained without having them linked to traditional emotions (Russell, 2009); it has been highlighted that core affect, and especially its second dimension "arousal" are connected to ANS-activities (Russell, 2003; Russell 2009). Core affect is viewed as a state of the central nervous system by Russell (2009). While Russell explains that specific emotions are not well relatable to somatic activity, he explicitly states that core affect is related to ANS-activity but yet highlights that it is a distinct concept from it (2003; 2009). He elaborates that arousal is determined at least in part by ANS-activity (2009). Because of these findings assessing the ANS-activities of a person appears to be the most suitable objective measure of core affect.

One of the most commonly applied measures of ANS-activities (more precisely the sympathetic nervous system) is electrodermal activity (Boucsein, 2012). This refers to a variety of electrical skin conductance measurements (Boucsein, 2012) in which the conductance is dependent on the activities of the endocrine sweat glands, in a way in which a higher presence of sweat (as caused by activity in the sudomotor nerves increase the conductance level of the skin (Alexander et al., 2005). Skin conductance measures can be further divided into their tonic form (skin conductance level) which provides gradual information of the level of skin conductance; and into phasic measures of skin conductance (skin conductance response), which is most suited for momentary measures, typically in response to an external stimulus (Boucsein, 2012; Alexander et al., 2005). An aspect that makes skin conductance measures especially suitable for the study at hand is the fact that skin conductance is often considered to be a measure of arousal which represents the second dimension of core affect (Alexander et al., 2005, Russell, 2009).

While not many studies have been conducted, which implement skin conductance estimates to investigated the relationship between core affect and somatic states, the few studies that conduct research in this domain usually apply skin conductance response (SCR) to measure the individuals arousal in response to experimental stimuli in a controlled environment (see for example Västfjäll & Gärling, 2007). While this kind of research certainly aims at conducting relevant investigations in the relationship between emotions and bodily states, the result may not be generalizable to everyday-life situations, as humans usually do not interact in controlled environments. Thus, also the estimates for correlation between skin conductance may not translate into the actual world. Thus, there is a need for an investigation into the relationship between ANS-activity and perceived emotional states in the context of the daily activities of individuals.

Intra-individual measurements

Previous studies into the correlation between emotions and bodily symptoms of emotions often suffered from issues of generalizability or from lack of a proper emotional theory. For example, Myrtek, Aschenbrenner, and Brügner (2005) did an experiment in which they aimed at testing the test subject's ability to discriminate between conditions with additional heart rate and random conditions without additional heart rate. They found that individuals did not differ in quality or frequency of the emotions they have reported in the two conditions. This study, however, made use of the basic emotional concepts rather than using a specific theory in order to provide a clear framework for defining emotions. This is as stated above not ideal since it has been acknowledged that traditional emotion concepts lack a proper universal definition and thus might get interpreted differently among different test participants. A more recent study investigated the correlation between ANS-measures and the state of core affect of the individuals by different measures (Västfjäll & Gärling, 2007). Despite potential problems with the generalizability of this study to everyday life, due to the fact that it was a laboratory study; another limitation of this paper was that the correlations were estimated on group level. Yet emotions have been found to be determined to a great extent by within individual variety (Russell, 2009). To counteract this, it is advisable to investigate the correlation between ANS processes and core affect intra-individually. In line with this, a study has indicated that for psychological processes that underlie individual variability should be analysed intra-individually, as the variety in intra-individual estimates tends to be substantially larger than in group-based estimates (Fisher, Medaglia, & Jeronimus, 2018). Another factor that supports this course of action, is the fact that also situational factors also influence the relationship between ANS and emotions (Russell, 2009) by for example altering the interpretation of the somatic symptoms depending on the context in which the individual is in. The test subjects in the study at hand might be exposed to different external factors from another during the course of this study. Furthermore, group-based analysis can provide limited meaningful data for the individuals because ecological group correlations give rise to the risk of the occurrence of ecological fallacies and thus (Piantadosi, Byar, & Green, 1988).

Goal of the paper

In order to counteract the limitations of traditional research on emotions and to work with a framework that provides clear states for the emotional research, Russell's theory of core affect will be used for this paper (Russell, 2009). Furthermore, because the continuous nature of core affect as described by Russell, this paper is investigating the relationship between somatic states and core affect over an extended time period of seven days, rather than investigating the relationship in the presence of sudden stimuli exposure as it has been done in previous studies (Västfjäll & Gärling, 2007). As stated before, this is also done because the repetitive application of devices that draw attention towards the own emotions among the test subjects are expected to enhance the overall awareness towards the own emotions over time (Farrell & Shaw, 1994). Thus, the outcomes from repeated measures over a seven-day period may result in more accurate data than a comparable study that applies a shorter time frame. Furthermore, because the importance of outside influences, and because of the desire to establish research that is useful for the medical sector the study at hand is conducted in the form of field research. The present study, therefore, aims at investigating the relationship between skin conductance and self-reported emotions during the daily activities of the participants. Due to the fact that this is a field study, SCL (and not SCR) is the most suited estimate of electrodermal activity, as participants are not expected to respond to a certain stimulus.

In summary, this paper investigates correlations between the self-perceived states of the participants in terms of Russell's conceptions of the Core affect, which will be applied in the form of a two-dimensional coordinate system; and the activity of the autonomic nervous system as measured by skin-conductance levels. This will be done in the form of a longitudinal field study.

Research question 1: *Are there correlations between the two dimensions of self-reported core affect and the measurement of ANS-activity based on skin conductance level estimates when analysing the data intra-individually?*

Based on the fact that a relationship between Russell's concept of core affect and ANS-processes has already been acknowledged and especially because the second dimension of core affect is arousal which in particular has been found to be related to ANS-activities (Russell, 2003; Russell 2009), it is assumed that a correlation between the self-reported core affect and ANS-activity as measured by skin conductance will be observed.

Furthermore, an investigation of potential observed differences between an intra-individual correlation analysis and an inter-individual correlation analysis of the data will be made. By this, an investigation is intended to determine the extent to which such parameters (i.e. the correlation estimates between core affect and skin conductance) can differ depending on the method of analysis.

Research question 2: *Are the group-level correlation estimates representative for the intra-individual correlations?*

Because of the fact that perception of emotion is determined by individual variability to a great extent and also influenced by situational factors, it will be assumed that correlations at the group level do not represent the correlation between SCL and core affect on the individual level. It is reasonable to assume that the correlation will differ to large extent between different participants due to the factors explained above. This potential diversity in correlations cannot be captured by a single combined correlation estimate.

Method

Design

This study applied longitudinal experience sampling, based on a within-subjects correlation design; aiming at gathering the variables of self-perceived core affect and ANS-activation as measured by the skin-conductance level of the participants in forms of quantitative field research.

Participants

The data for this paper is drawn from two sources. For one the data from 16 participants was used from an already existing data set gathered for a previous study. Additionally, 11 participants were recruited just for the study at hand. Concerning the demographic characteristics of the data set from the previous study, it can be stated that the mean age was 30.63 (SD = 12.96), ages ranging from 21 to 70. Of this data set, 5 participants are identified as female and 11 are identified as male. 13 of the participants were of German nationality and 3 of the participants were of Dutch nationality.

Concerning the demographics of the data of the participants that were recruited for the current study, the mean age was 27.27 (SD = 13.82) and that the ages ranged from 18 to 55. This data set included eight females and three males, and all of these participants had German citizenship. These characteristics give the overall combined data set a number of 24 Germans and three Dutchman; of which are 14 males and 13 females. For the data collection that was exclusively part of the present study, a convenience sample was applied with two inclusion criteria; namely the participants had to be above the age of 18 and had to be in the possession of a smartphone or a tablet which was necessary in order to run the “The incredible intervention

Machine"-app that was required to be used in this study. Furthermore, some of the participants received a reward in form of course credits on the online platform SONA, which is a credit system of the University of Twente, that rewards students for participation in ongoing studies (Libbertz, 2019). Besides that, no other rewards have been offered to the study participants. Before the data for this paper was gathered the ethics committee of the BMS faculty of the University of Twente gave approval for this study.

Materials

This study is based on the use of two main devices: namely the Empatica E4 which is a device that can monitor different measures of ANS activity like skin conductance levels (Empatica, 2019); and the "The incredible intervention machine"-app (TIIM) which was used for this paper as a means to provide the participants with a quadrant that represented dimensions of core affect (valence and arousal) in a way that the participants were able to locate their current self-perceived state in the quadrant.

The Incredible Intervention Machine.

TIIM is a software that was created by the BMSLab at the University of Twente (Appadvice, n.d.; University of Twente, n.d.) that can be run on Android and iOS and was initially designed to implement interventions (University of Twente, n.d.); but can also be used to ask questions to study participants in predefined time intervals. The version that was used in this study was 1.3.7, which is able to run on Android versions 4.1 or higher (Google Play, n.d.) and iOS version 9.0 or higher (App Store, n.d.). Thereby the main reason to apply this technology was that it runs on mobile devices such that the participants could easily carry the required devices around during their daily routines. In the app, a visualization of the conceptions of core affect in the form of a coordinate system was displayed based on Russell's conceptions thereof (2009). The *y*-axis symbolised the arousal dimension of core affect, ranging from a low energy level on its lower end to a high energy level on its upper end. Subsequently, the *x*-axis symbolised the valence dimension of core affect, ranging from an unpleasant feeling on the lower end to a pleasant feeling on its upper end (Figure 1). The questions posed by the TIIM-app as well as the coordinate system ("*core quadrant*") were translated into English, and German in order to enable the inclusion of participants of different languages.

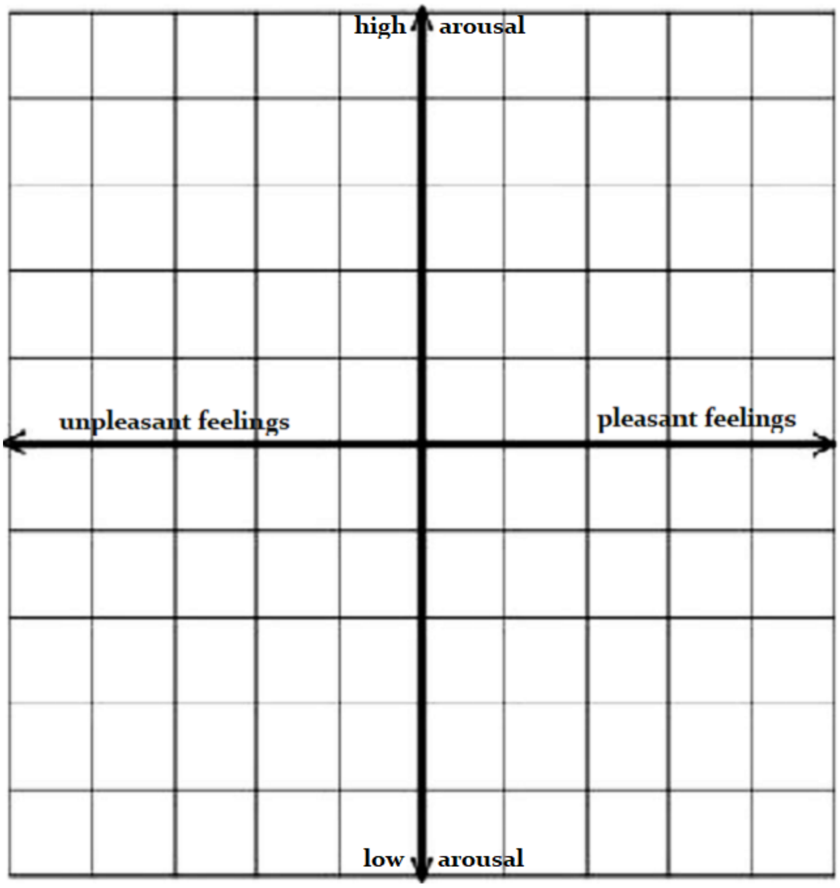


Figure 1. English version of the core affect based on Russell (2009).

Empatica E4.

The E4 is a wristband-like device, developed by the BMSLab of the University of Twente, which is placed just above the wrist joint and with which among other measures like heart rate, electrical conductance levels of the skin can be captured by means of two inbuilt electrodes on the inner side of the device. (Empatica, n.d.; Empatica, 2019). As stated above the measurement of interest in this study is tonic SCL, which can be measured in units of microSiemens (μS) (Empatica, 2019). A previous study has already established the *usability* and the *wearing comfort* of this exact device, using a 5-point Likert scale of ten items based on the System Usability Scale (SUS) for the former, and a semi-structured interview for the latter (Lier et al., 2017). This might be important as these factors may influence the perceived emotional states of the participants in the study at hand. The study indicated that most participants perceived the E4 as easy to use and felt that no additional technical support was needed for the device despite reporting a certain level of frustration from the use of the device. Furthermore, the participant in that study rated the wearing comfort of the E4 as neutral to positive and all the participants indicated that the process of getting used to the device was fast and with ease. A downside that participants frequently reported, however, was that the wristband is not waterproof which was perceived as “burdensome” (Lier et al., 2017). Due to the fact that the overall perception of the E4 was positive to neutral, the devices are not expected to have a severe influence on the outcomes of this study and thus seems to be an appropriate tool to gather information on the skin conductance level of the study participants.

Procedure

The data collection for the data sample that was gathered specifically for this study occurred in April and May of the year 2019, and all participants were asked to give informed consent just before the data collection, followed by a measurement period of seven days per participant. Furthermore, it was highlighted to them that the study at hand is not intended to be a medical study of emotions and the right to withdraw at any point in time was emphasised as well. To each participant, an instruction email was sent that introduced them to the steps that have to be taken in order to run the data collection process properly. Furthermore, for each participant a meeting was arranged in a non-controlled setting (usually at their private home), in which the functionalities of the TIIM app as well as of the E4 were explained to them. During these meetings, the researchers provided their contact information to the participants in order to enable them to ask questions in case of confusion or problems. The researchers guided the participants through the creation of a TIIM account as well as an Empatica account which were both required to capture the data of each individual. The entire data collection took place in non-controlled environments (I.e. during the daily lives of the test subjects).

For this study, the TIIM app was set up in such a way that the participants had to indicate their self-reported core affect on the quadrant. Thereby, participants had to determine their position on the quadrant concerning the last minute, as well as concerning the last two hours. These two questions were asked in an interval of two hours, starting either at a predefined point in time or at the moment that the participant was introduced to the software, depending on the participant's preferences. Thereby the questions were asked at day time as well as night time. In order to remind them, a notification was placed as soon as the two questions were available to be answered, and if no response was recorded within 15 minutes a second notification was sent. If, however within 30 minutes no response was given by the respondent, the question of this measurement point was marked as unanswered and the participant was able to answer the next question after the predefined time interval has surpassed. This procedure was repeated over a time frame of seven days, in which the two questions were asked each 12 times a day and 84 times in total.

At the beginning of the experiment, the participants were introduced to the functionalities of the E4 and were instructed to synchronize the data from the E4 to their Empatica-accounts on a semi-daily basis. It was also explained to the participants that they could not wear the E4 in all situations due to the limitations of the technology (for example not while showering or taking a bath). During the data collection process, the participants were required to wear the E4 during daytime for 7 days while simultaneously filling out the questionnaire provided by the TIIM-app. Thereby the E4 measured the skin conductance level with a default sampling frequency of 4 Hz (Empatica, 2015). After the seven-day period had passed, the participants were instructed to download the data from their Empatrica-accounts and to hand it over to the researchers alongside the E4 devices themselves. Additionally, the demographic characteristics gender, age, and nationality have been gathered from the participants and before the data analysis was conducted, all data sets have been anonymised, by reducing the participant information to the aforementioned characteristics.

Data Analysis

Two kinds of information were intended to be gathered from the participants (alongside some demographic information): self-perceived (emotional) state in terms of Russell's conceptions of core affect, and the skin conductance level of the participants as a means to access their ANS-activity. The entire study took place under natural conditions and no confounding variables were monitored.

The data on the skin conductance that was gathered by the E4 devices was processed by means of the EDAexplorer (Taylor et al., 2015). Thereby, Python was used to run the EDAexplorer and to synchronize the time stamps of the core affect data with the timestamps of the SCL and for other measurements gathered by the E4. Thereby mean SCL, as well as the summed peak and amplitude scores, could have been created for the one-minute time frame and for the two-hour time frame. The trough-to-peak analysis that was applied made use of a minimum SCR amplitude (not further used in the present study) of .001 µSiemens.

Afterwards, the data set has been screened for participants that have less than five measurement points. The data for these participants will not be included. All statistical analysis was conducted with SPSS 24. The standard deviation and the mean of the age of the participants were calculated, followed by a descriptive statistical analysis of the variables in question, and a Shapiro-Wilk test to control for normality. The latter was assisted by a histogram that displays the normal distribution of the variables.

Hereafter, the data was found not to be normally distributed a Spearman's rank-order correlation was conducted first on the intra-individual level, to test the first hypothesis, and then on group-level to test the second hypothesis. These correlations have been analysed for the two variables valence and arousal (as retrieved from the core quadrant) each in relation to SCL measure. This has been done for the two hours and the one-minute mark. Additionally, these correlations were assisted by histograms that display the distribution of the correlations.

Results

Of the data set containing 26 participants, the data of two participants was excluded. For one of them (participant 47) only one SCL-measurement was available and because for the other one (participant 42), only four measurement points were available. Leading to a dataset of 24 participants.

The dataset contains six measurement variables per participant. The mean SCL measured over a one-minute period before the proposition of the core affect question to the participant, and the self-perceived arousal and valence level driven from the indication by the participant on the core quadrant. As well as the counterparts of these three measurements for the two-hour period before the proposition of the question. The mean of the SCL level for one minute was 1.22 μS (*SD* = 10.81) while the mean measured for the last two hours before the assessment of the core affect was 1.21 μS (*SD* = 6.82). The level of arousal that the test subjects indicated over a one-minute period was 4.1 (*SD* = 44.75) and the mean arousal for two hours prior to the measurement point was 4.14 (*SD* = 45.59). Valence, as indicated for the past minute, had a mean of 6.61 (*SD* = 45.585) and 5.77 as indicated for the past two hours (*SD* = 45.872). Furthermore, the SCL scores per participant have been visualised for the one-minute time frame and for the two-hour time frame, in the form of box plots (Figure 2; Figure 3).

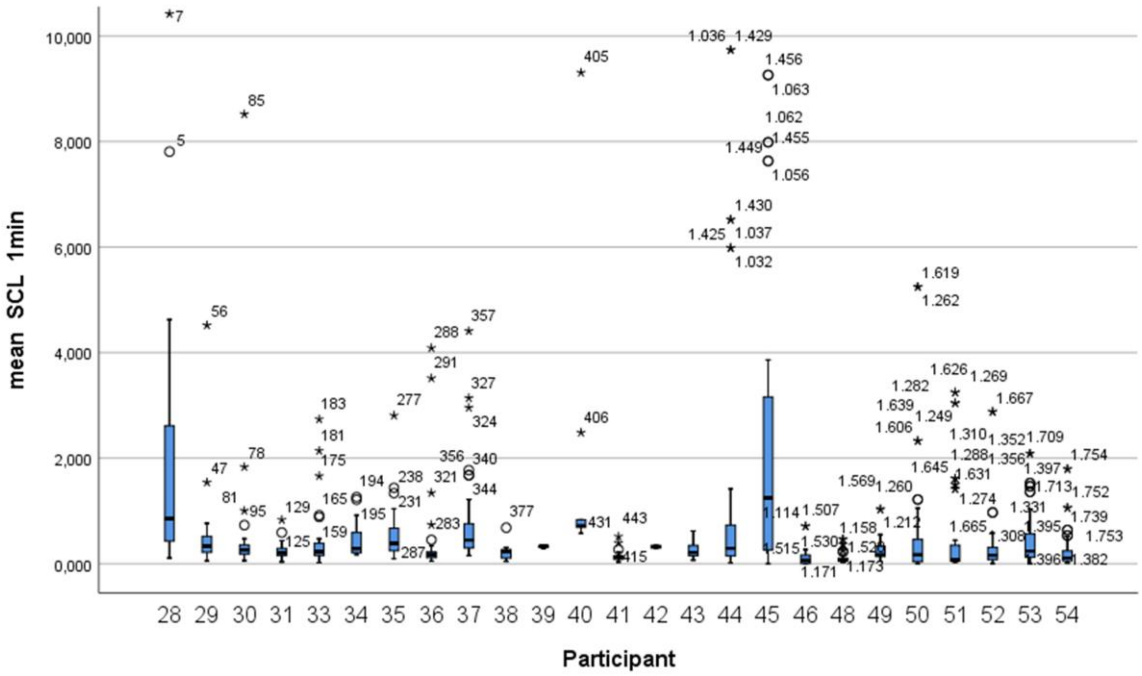


Figure 2. Boxplot of the SCL scores per participant for the one-minute interval.

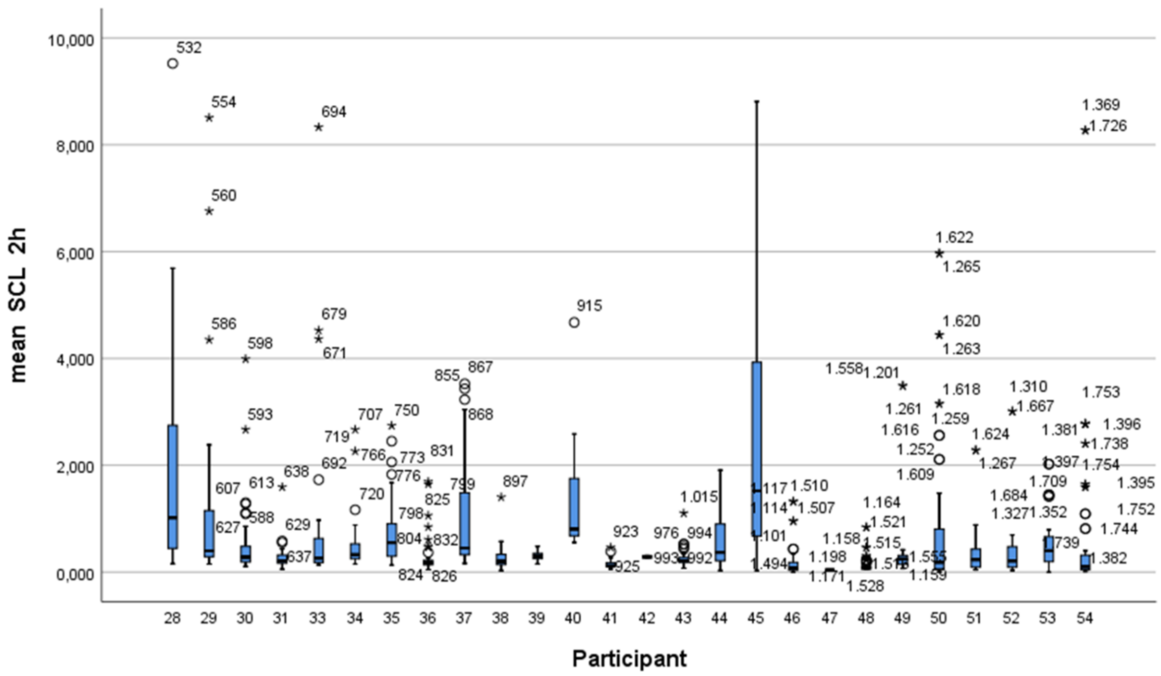


Figure 3. Boxplot of the SCL scores per participant for the two-hour interval.

The results of the Shapiro-Wilk test indicate that the data is not normally distributed (Sig < 0.01) for all variables in question (Appendix A). Additionally, histograms were created for the two-hour interval which displays the normal curve of the variables in question. This was done group-based as well as intra-individually. Due to simplicity reasons however, in this paper only the histogram for the group-based analysis will be displayed (Appendix B, C & D).

To answer the first research question, the data for each participant was analysed individually. Thereby the results will be reported in two segments, first addressing the correlation of self-reported feelings concerning the last minute with the corresponding SCL-scores; thereafter, addressing the correlation concerning the self-reported feelings regarding alongside the corresponding SCL-scores for the last two hours.

For the past minute prior to the measurement point, the arousal measurements of two participants correlated significantly. For one there was a significantly negative relationship of arousal with SCL (r_s = -.733; p < .05), and for one there was a significantly positive relationship (r_s = .469; p = .05). All the correlation estimates for arousal and SCL in the one-minute interval are also displayed in a histogram (figure 4).

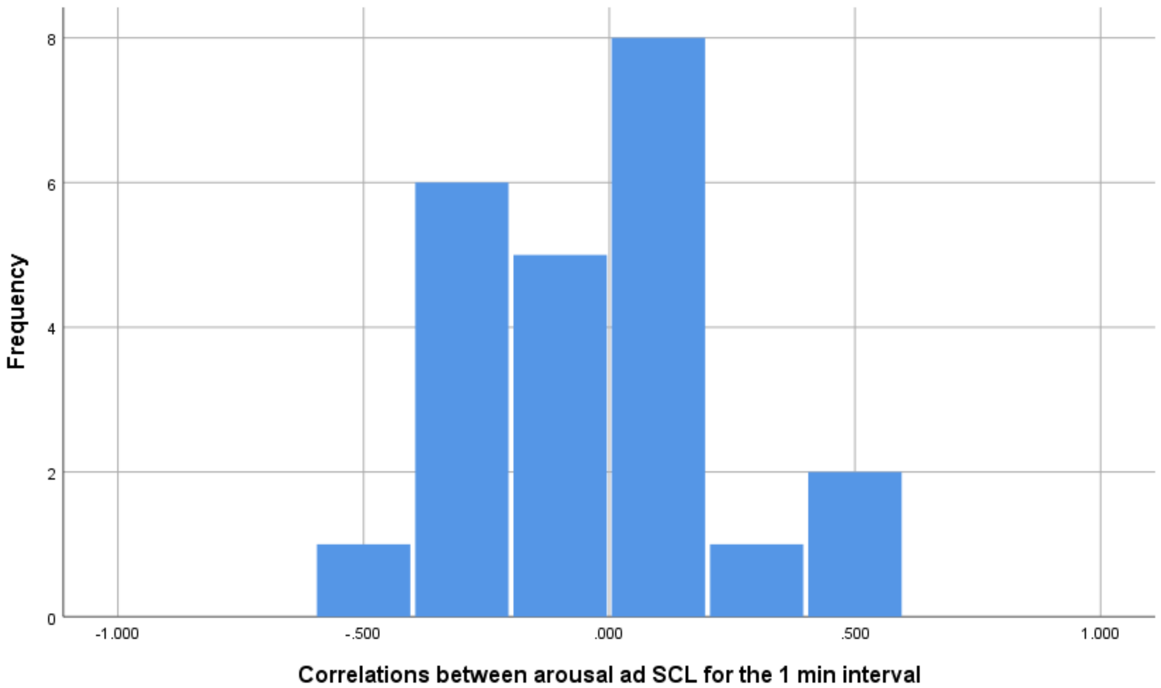


Figure 4. Histogram with normal curve for correlations between SCL and arousal for the one-minute interval.

The spearman correlation of valence regarding the one-minute interval indicated 7 significant correlations. Two of these correlations where positive: participant 37 (r_s = .340; p < .05) and participant 40 (r_s = .644; p < .05). Furthermore, five of these correlations were negative: participant 30 (r_s = -.372; p < .05), participant 31 (r_s = -.485; p < .05), participant 43 (r_s = -.368; p < .05), participant 44 (r_s = -.375; p < .05) and participant 46 (r_s = -.342; p < .05). The correlations for valence and SCL are displayed in a histogram (figure 5). In summary 17 significant correlations have been found in both time frames taken together. Seven of these correlations were positive and ten negative.

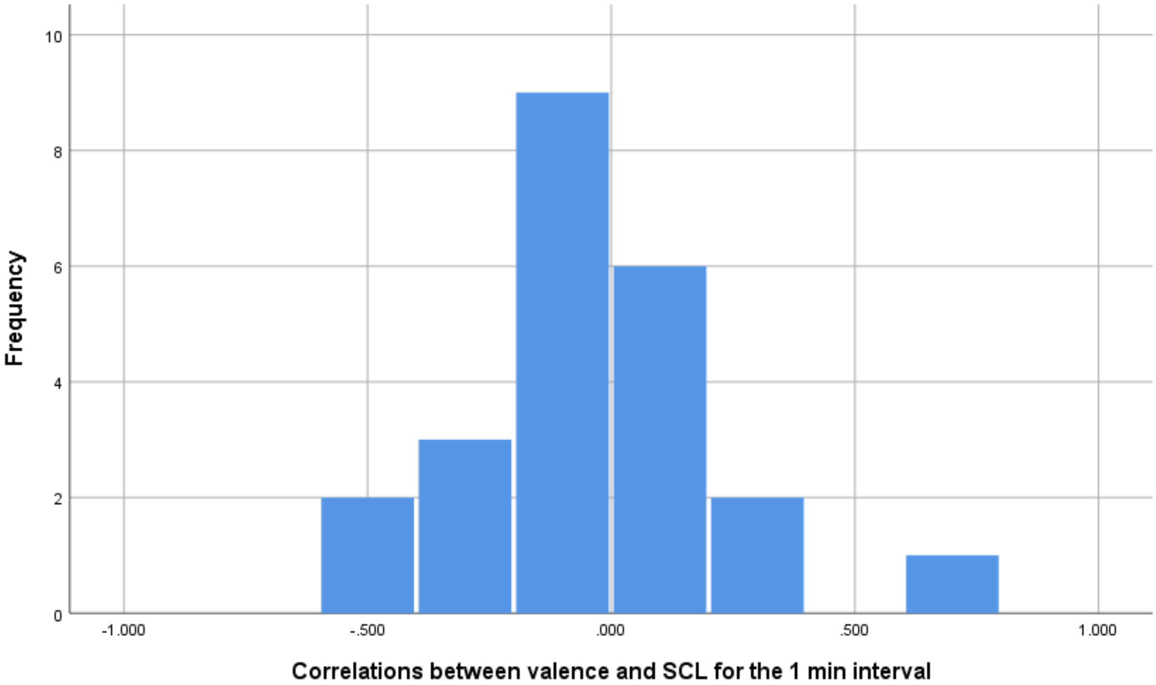


Figure 5. Histogram with normal curve for correlations between SCL and valence for the one-minute interval.

Regarding the two-hour time frame, four significant correlations have been found between arousal and SCL. Three of these correlations were negative: participant 28 (r_s = -.586; $p < .05$), participant 34 (r_s = -.779; $p < .05$) and participant 53 (r_s = -.411; $p < .05$); and one (participant 49) was positively correlated (r_s = .638; $p < .05$).

For the correlation of valence with SCL over the two-hour time frame, four significant correlations have been found as well. One of which (participant 28) was negative (r_s = -.385; $p < .05$) and three of which were positively correlated: participant 45 (r_s = .397; $p < .05$), participant 49 (r_s = .407, $p < .05$) and participant 50 (r_s = .316; $p < .05$). The correlations between arousal and SCL for the two-hour time frame are displayed in figure 6, while the correlations between valence and SCL are visualised in figure 7.

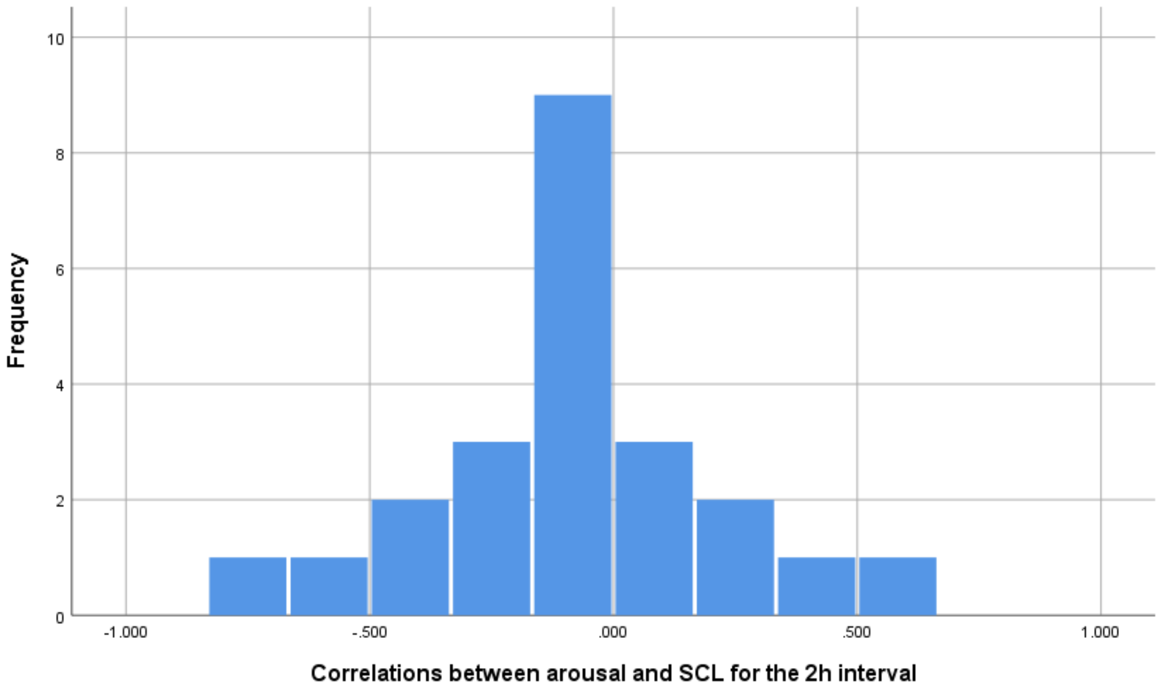


Figure 6. Histogram with normal curve for correlations between SCL and arousal for the two-hour interval.

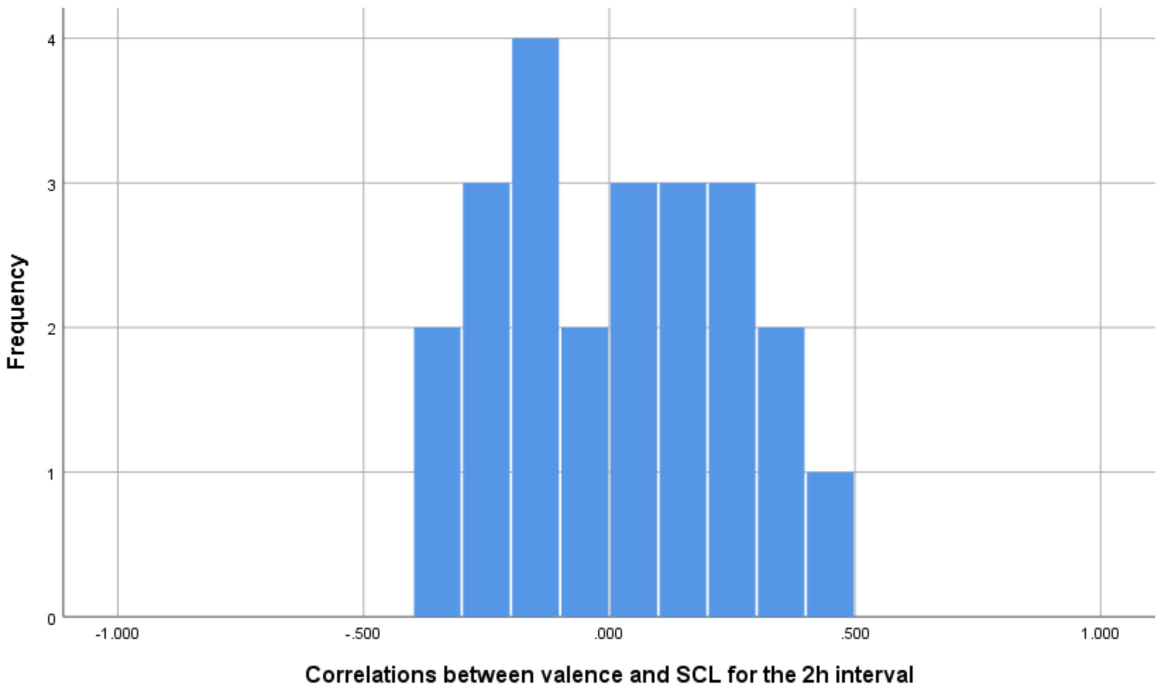


Figure 7. Histogram with normal curve for correlations between SCL and valence for the two-hour interval.

For the second research question, the per participant averages of the variables SCL, arousal, and valence have been calculated. Thus, the data has not been screwed by the number of measurement points, which is far greater than the number of participants. Thereafter, the correlations between the per participant averages of arousal and SCL have been correlated; as well as the per participant averages of valence and SCL. This has been done for the two-hour time frame as well as for the one-minute time frame with a two-tailed bivariate Spearman's rank-order correlation. The SCL measurements for the entire sample did not correlate significantly with the indicated arousal of the entire sample for the one minute time frame (r_s = -.114; p = .596) (Appendix E). Also, the Spearman correlation analysis for SCL and valence for the one-minute time frame indicated a non-significant negative correlation (r_s = -.217; p = .307) (Appendix F). The Spearman correlation for the two-hour time frame indicted an insignificantly negative correlation between SCL and arousal (r_s = -.223; p = .259) (Appendix G) and an insignificant negative correlation between SCL and valence (r_s = -.371; p = .074) (Appendix H). Thus, it can be noted that none of the correlations on the inter-individual level were significant.

Discussion

This study was aimed at investigating the relationship between SCL and perceived states of mood intra-individually and at contrasting the outcomes of this kind of analysis with the inter-individual correlation estimates. In summary, the outcomes of the various analyses indicate that, against the expectations, little statistical evidence was present that suggest prominent intra-individual correlations between SCL-measurements and the two dimensions of core affect. To the contrary, SCL-measures correlated insignificantly negative with arousal and valence in the inter-individual analysis for the one-minute time frame as well as the two-hour time frame. When reflecting the findings of the first part of the analysis on the first hypothesis (i.e. the intra-individual analysis), it can be stated that even though some correlations were present, more often than not SCL does not correlate with the two dimensions of core affect on an intra-individual level. The most correlations have been found, however, between SCL and valence in the one-minute time frame. The outcomes of the second part of the analysis support the second hypothesis, which stated that the correlations at the group-level are not representative of intra-individual correlations. This conclusion is based on two observations. First of all, even though correlations were uncommon among the individual level, on the group level the correlations were weak but clearly significant. It is noteworthy, that while for the intra-individual analysis seven of the 17 correlations (as found in both time intervals) have been found to be positive, no significant correlation was found on the inter-individual level. Moreover, the correlation estimates that were measured on the group level between SCL and the two dimensions of core affect fails to indicate the multifaceted nature of correlations on the individual level, which is characterised by positive and negative correlations. Nonetheless, most correlations that were present in the intra-individual analysis were also negative just as in the group-level analysis (which was however not significant).

Theoretical reflection and implications

Because SCL is, as stated above, a measure of ANS-activity, the lack of correlation that has been found in the group level analysis seems to contradict Russell's (2003) conception of arousal as being interconnected with the central nervous system. Furthermore, even if the correlations on the group level were significant, all correlations are yet negative (just like most correlations on the intra-individual level). This seems for one counterintuitive as one would expect a positive correlation between arousal and emotional states, and also is in contrast with previous research (Västfjäll & Gärling, 2007). This is especially salient when reflecting on the fact arousal is viewed as a state of the central nervous system (Russell, 2003). Russell did note than self-reported arousal might be only weakly correlated to autonomic measures of arousal (Russell, 2003). Yet, the aforementioned study that investigated the relationship between core affect and skin conductance on group level has found strong positive correlation between SCR and core-affect (Västfjäll & Gärling, 2007). The difference in the strength of the association may come from the fact that in the present study the subjects interacted in a normal daily environment, rather than reacting to a stimulus in an isolated laboratory setting. This might cause differences because factors such as the social setting and so on have not been recorded and thus cannot be controlled for in the analysis unlike in a laboratory study. Furthermore, only a linear relationship between somatic states and self-reported emotions has been tested. The outcomes of the present paper, therefore, seem to support the conception that self-perceived emotions are influenced to a large extent by situational factors (Russell, 2009) such as the precise social setting or peer pressure and even the anticipation of events by the individual (Russell, 2009). This is because in isolation of such factors (i.e. in laboratory settings) the correlation between core affect and ANS-measures has been found to be positive and strong. But further research is required to support this assumption. A study by Myrtek, Aschenbrenner, and Brügner is supporting the low estimates for correlation between SCL and core affect that were present in the study at hand, as it found that physiological activation plays only a secondary role in the perception of emotions (2005). Cognitive schemata and personality dimensions have been found to be more important in this process.

Also, the observation that at the intra-individual level some correlations (seven of the 17) were found to be positive while the group-level analysis yielded in exclusively non-significant negative correlations requires further reflection. First of all, it can be assumed that an ecological fallacy (Piantadosi, Byar, & Green, 1988) such as a Simpson's paradox did not occur. In a Simpson's paradox, the correlation on the inter-individual level would display a reversed sign of the association to the intra-individual analysis or would lack an association altogether (Tu, Gunnell, & Gilthorpe, 2008). While the latter cannot be ruled out due to the insignificance of the correlations on the inter-individual level, the strength and sign of the (insignificant) group-level correlations is largely in line with the individual estimates. The differences in the significant likely arise from the fact that each study participant just equates to one data point per variable in the inter-individual analysis, while on the intra-individual analysis, a lot of participants show a larger amount of measurement points than the actual sample size. This, of course, influences the significance of the correlation. Besides that, the contrast between the low presence of correlations in the intra-individual analysis and the insignificant correlation outcomes on the inter-individual analysis leads one to assume that in the study at hand there was no prominent issue with the generalizability of group-level outcomes to individual outcomes. This is an important conclusion to report as Fisher, Medaglia and Jeronimus (2018) found correlations of psychological processes to often be not generalizable from inter-individual level to the intra-individual level. This is due to common variability on the individual level and to time variability (Fisher, Medaglia, & Jeronimus, 2018). The high level of face validity of generalizability from the inter-individual level to the intra-individual level is furthermore interesting as emotions are known to be dependent on individual factors (such as thought processes or past experience) (Russell, 2009), and thus may be vulnerable to low group-to-individual generalizability.

Another noteworthy observation from the present study is that SCL seems not to be universally better predicted by arousal than by valence. In the intra-individual analysis, the correlation estimates of arousal and SCL tended to be stronger, yet for example, in the one-minute interval, more correlation between valence and SCL have been found than between arousal and SCL. Moreover, among the inter-individual correlation estimates, none were significant, however, even if we were to ignore the lack of significance, the correlations between valence and SCL still would be stronger than those between SCL and arousal. On face value, it does not appear to be a worse predictor of SCL than arousal is. This is a particular conclusion as in a similar study, Västfjäll and Gärling investigated the correlation of SCR, with a set of different measures of the theory of core affect by Russell on a group-level and found a higher correlation between arousal and SCR than between valence and SCL (2007). This finding is important as SCL and SCR are both measures of electrodermal activity (Boucsein, 2012) and arousal is assumed to be a state of the ANS (Russell, 2003; Russell, 2009). However, while Västfjäll and Gärling found a strong positive correlation between arousal and SCR (r = .88) on an interindividual analysis, the results of the present paper indicate a weak negative correlation on the intra-individual level at least. Also, the correlation between valence and SCR was positive in their paper (r = .42) unlike in the paper at hand. As stated before, such an observed difference may be due to the difference in the conduction of the experiment when compared to the study at hand. More precisely, in the study at hand, a lot of confounding variables may have influenced the association between the variables (especially regarding social settings that influence the interpretation of emotions), unlike in a laboratory setting.

It is furthermore plausible to assume that over a longer period of measurement, more significant correlations would have been found in the intra-individual analysis, because, as Farrell and Shaw (1994) have argued, the regular employment of technologies that lead individuals to increase their awareness of the personal emotional states, like the TIIM-app used in this study, can increase the overall emotional awareness over time. Such data might be especially useful in medical settings in which individuals have problems monitoring or expressing their own emotions, but maybe less so useful in regular daily interactions as individuals are usually not overtly aware of their own emotions. This prediction should, however, be validated by further research.

In summary, it can be stated that a weak negative correlation between SCL and each dimension of core affect is the dominant form of correlation on the intra-individual level. The group-analysis, while capturing the general trend of the association in terms sign, fails to display the multi-faced nature of the relationship between core affect and SCL and lacked significance. It can also be assumed that differences when comparing the outcomes to other studies perhaps can be attributed to situational and individual factors. Nonetheless, the exact nature of this relationship remains unclear. Therefore, further research is needed in order to investigate the particularities of the influence of ANS-activity on core affect and the interference by situational and personal factors.

Strong and weak points of the study

A strong point of this study is that it is, to the knowledge of the author the first one which aims at investigating the relationship between core affect and ANS-activity intra-individually in daily life. As stated before, this approach excludes the occurrence of a Simpson's paradox that may appear in group analyses. Furthermore, this study made use of measurement devices that are perceived to have a high wearing comfort and ease of use, which limits the potential influence of the technology on the participant's mood (Lier et al., 2017). This approach of experience sampling is furthermore assumed to be a suitable tool for identifying variability over time as well as situational and personal factors (Myin-Germeys et al., 2018). This is especially important as the perception of emotions is influenced by this kind of determinants (Russell, 2009).

Also, the fact that the study was not conducted under laboratory conditions suggests high generalizability to everyday life when compared to studies under controlled conditions. As this field is of potential medical interest, for example when dealing with patients who suffer from communication disabilities (Picard, 2009; Noordzij & Laroy-Noordzij, 2014) personality disorders (Derks, De Visser, Bohlmeijer, & Noordzij, 2017) or who possess personality traits such as Alexithymia (Valdespino, Antezana, Ghane, & Richey, 2017); generalizability to daily situations is an important aspect of the research.

A limitation of this study is, as it is part of a bachelor thesis, restricted time available to conduct the research. Therefore, only one measure of ANS-activity (namely SCL) was used as a means to investigate the correlation between core affect and ANS-activity in general. This is problematic because as Russell points out core affect should not be limited to a single measurement (2003). However, similar measures of skin conductance have been shown to be related to both dimensions of core affect, and especially strongly related to the arousal dimension (Västfjäll & Gärling, 2007); even though a multi-measure approach might have been preferable. Also, the general approach of experience sampling methodology was well suited for the measure in question due to the variable nature of emotion in individuals (Russell, 2009; Myin-Germeys et al., 2018).

Another limitation that resulted from time constraints is that the data in this study was collected in a seven-day period, which may not be ideal for a complex subject like emotion, as emotional awareness increases over time when awareness increasing technologies (such as the TIMM app) are regularly employed (Farrell & Shaw, 1994). A lot of correlations on the intra-individual analysis were not significant, therefore this study should be replicated to ensure that low presence of significant correlations in the intra-individual analysis are not just due to an issue of power. Nonetheless, it can be said that in this study during these seven days a lot of measurement points have been employed which enlarges the data set.

Suggestions for further research

There is a need for a long-term study investigating the correlation between ANS-measures intra-individually. Thereby special attention should be drawn towards individual and situational factors that influence the relationship between ANS-activity and the two dimensions of core affect. This is due to several reasons. A long-term employment of technologies that increase emotional awareness is theorized to increase emotional awareness in general. Thus, it can be assumed that study participants become better able to monitor and thus ultimately report their emotional states more accurately in a long-term study (Farrell & Shaw, 1994). Moreover, there is a general lack of emotional research, and research on core affect, in particular, that takes an intra-individual approach. Research on the individual and situational factors of emotional perception is important. By this, it could be possible to better understand the relationship between SCL and core affect and also understand how different studies can yield in very different correlation estimates between somatic measures and core affect, or somatic measures and emotions in general.

References

Alexander, D., Trengove, C., Johnston, P., Cooper, T., August, J., & Gordon, E. (2005). Separating individual skin conductance responses in a short interstimulus-interval paradigm. *Journal of Neuroscience Methods, 146*(1), 116-123. doi:10.1016/j.jneumeth.2005.02.001

Appadvice. (n.d.). The Incredible Intervention Machine created by The University of Twente. Retrieved May 6, 2019, from https://appadvice.com/app/tiim/1229896853

App Store. (n.d.). TIIM. Retrieved May 6, 2019, from https://itunes.apple.com/nl/app/tiim/id1229896853?mt=8

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). New York: Springer.

Brouwer, A., Beurden, M. V., Nijboer, L., Derikx, L., Binsch, O., Gjaltema, C., & Noordzij, M. (2018). A Comparison of Different Electrodermal Variables in Response to an Acute Social Stressor. *Symbiotic Interaction Lecture Notes in Computer Science,* 7-17. doi:10.1007/978-3-319-91593-7_2

de Bruin, L., Derks, Y., & Noordzij, M. L. (2017). Sense-IT! Designing a therapists specific section of the app by the therapists mental model of emotion and an user centered design model. University of Twente. Retrieved from: https://essay.utwente.nl/73439/1/deBruin_MA_HFE.pdf

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205-218. doi:10.1037//0022-3514.70.2.205

Derks, P. M. J., De Visser, T., Bohlmeijer, E. T., & Noordzij, M. L. (2017). mHealth in Mental Health: How to efficiently and scientifically create an ambulatory biofeedback e-coaching app for patients with Borderline Personality Disorder. International Journal of Human Factors and Ergonomics, 5(1), 61–92.

Empatica. (2015). E4 wristband user's manual. Retrieved from: https://empatica.app.box.com/v/E4-User-Manual

Empatica. (2019, January 29). What should I know to use EDA data in my experiment? Retrieved May 2, 2019, from https://support.empatica.com/hc/en-us/articles/203621955-What-should-I-know-to-use-EDA-data-in-my-experiment-

Empatica. (n.d.). E4 wristband from empatica. Retrieved May 6, 2019, from https://www.empatica.com/research/e4/

Farrell, J., & Shaw, I. (1994). Emotional awareness training: A prerequisite to effective cognitive-behavioral treatment of borderline personality disorder. Cognitive and Behavioral Practice, 1(1), 71-91. doi:10.1016/S1077-7229(05)80087-2

Feldman Barrett, L. (2008). The Science of Emotion: What People Believe, What the Evidence Shows, and Where to Go From Here. In J. Blascovich & C. R. Hartel (Authors), *Human behavior in military contexts* (pp. 189-216). Washington, DC: National Academies Press

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*(3), 464-486. http://dx.doi.org/10.1037/0096-3445.113.3.464

Fernández-Dols, J., Wallbott, H., & Sanchez, F. (1991). Emotion category accessibility and the decoding of emotion from facial expression and context. *Journal of Nonverbal Behavior, 15*(2), 107-123. doi:10.1007/bf00998266

Fisher, A., Medaglia, J., & Jeronimus, B. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the United States of America*,115(27), 6115. doi:10.1073/pnas.1711978115

Google Play. (n.d.). TIIM. Retrieved May 6, 2019, from https://play.google.com/store/apps/details?id=nl.bmslab.utwente.tiimapp

Libbertz, L.B. (2019). *Can reactivity to experience sampling be used to decrease alexithymia?* (Unpublished master's thesis). University of Twente. Netherlands. Retrieved from https://essay.utwente.nl/77178/1/Libbertz_BA_PSY.pdf

Lier, H. G., Oberhagemann, M., Stroes, J. D., Enewoldsen, N. M., Pieterse, M. E., Schraagen, J. M., . . . Noordzij, M. L. (2017). Design Decisions for a Real Time, Alcohol Craving

Study Using Physio- and Psychological Measures. Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors Lecture Notes in Computer Science, 3-15. doi:10.1007/978-3-319-55134-0_1

Mulligan, K., & Scherer, K. R. (2012). Toward a Working Definition of Emotion. *Emotion Review, 4*(4), 345-357. doi:10.1177/1754073912445818

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry, 17*(2), 123-132. doi:10.1002/wps.20513

Myrtek, M., Aschenbrenner, E., & Brügner, G. (2005). Emotions in everyday life: An ambulatory monitoring study with female students. *Biological Psychology, 68*(3), 237-255. doi:10.1016/j.biopsycho.2004.06.001

Noordzij, M. L.; Dorrestijn, S. M.; Van den Berg, I. A. (2016). An idiographic study into the physiology and selfreported mental workload of learning to drive a car. In N. M. D. de Waard, K.A. Brookhuis, A. Toffetti, A. Stuiver, C. Weikert, D. Coelho, D. Manzey, A.B. Ünal, S. Röttger (Ed.), Proceedings of the Human Factors and Ergonomics Society (pp. 1–13). Groningen, The Netherlands.

Noordzij, M. L., & Laroy-Noordzij, M. E. (2014). MoodRadar-DAVID: providing caretakers with real-time, personalized insight in electrodermal activity of clients who show severely challenging behavior. In Human Factors –Europe Chapter . Lissabon, Portugal.

Piantadosi, S., Byar, D., & Green, S. (1988). The ecological fallacy. *American Journal of Epidemiology, 127*(5), 893-904. Retrieved March 22, 2019, from http://www.scientificintegrityinstitute.org/aje050188.pdf

Picard, R. W. (2009). Future affective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1535), 3575-3584. doi:10.1098/rstb.2009.0143

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145-172. doi:10.1037/0033-295x.110.1.145

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion, 23*(7), 1259-1283. doi:10.1080/02699930902809375

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015). Automatic identification of artifacts in electrodermal activity data. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. doi:10.1109/embc.2015.7318762

Thoits, P. A. (2010). Stress and Health: Major Findings and Policy Implications. *Journal of Health and Social Behavior, 51*(1_suppl). doi:10.1177/0022146510383499

Tu, Y., Gunnell, D., & Gilthorpe, M. S. (2008). Simpsons Paradox, Lords Paradox, and Suppression Effects are the same phenomenon – the reversal paradox. *Emerging Themes in Epidemiology, 5*(1). doi:10.1186/1742-7622-5-2

University of Twente. (n.d.). TIIM (The Incredible Intervention Machine). Retrieved May 6, 2019, from https://bmslab.utwente.nl/knowledgebase/tiim/

Valdespino, A., Antezana, L., Ghane, M., & Richey, J. A. (2017). Alexithymia as a Transdiagnostic Precursor to Empathy Abnormalities: The Functional Role of the Insula. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.02234

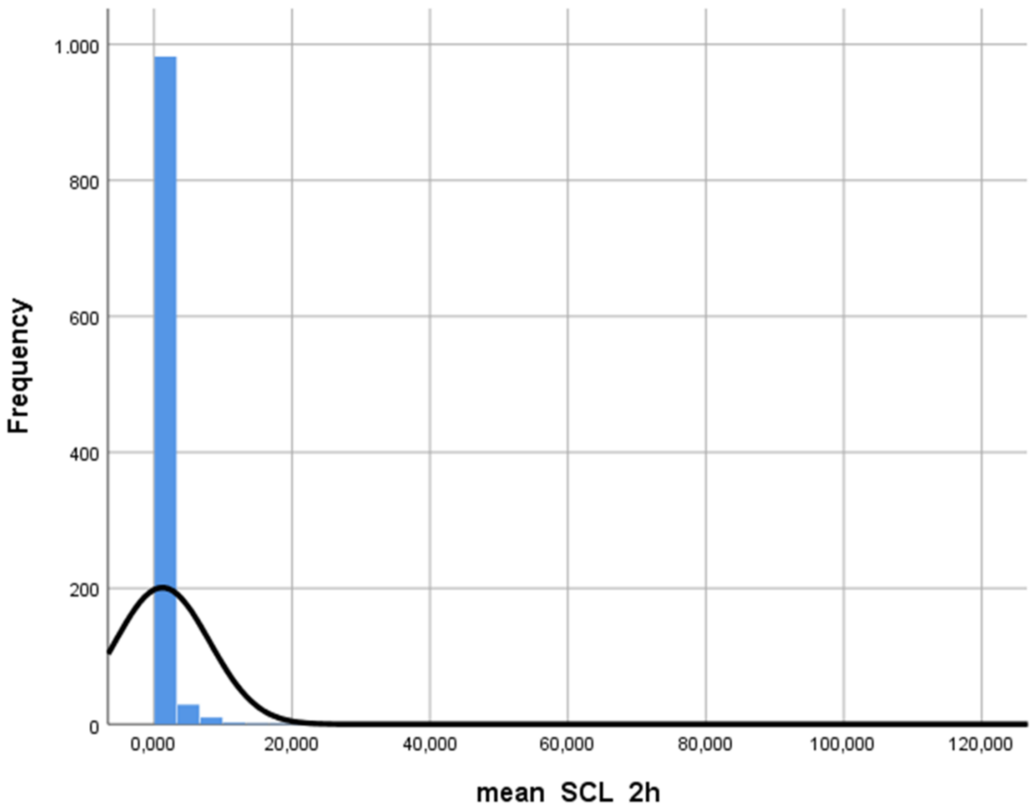
Västfjäll, D., & Gärling, T. (2007). Validation of a Swedish short self-report measure of core affect. *Scandinavian Journal of Psychology, 48*(3), 233-238. doi:10.1111/j.1467-9450.2007.00595.x

Appendix

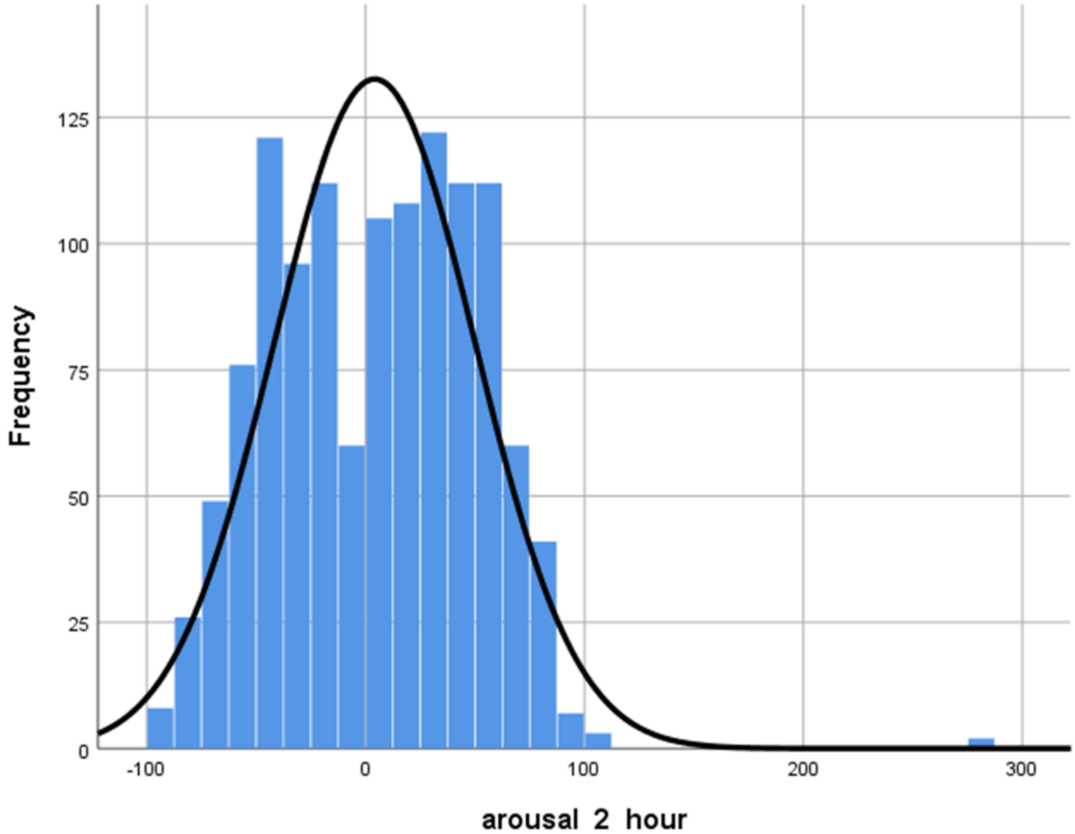
Appendix A

Tests of Normality

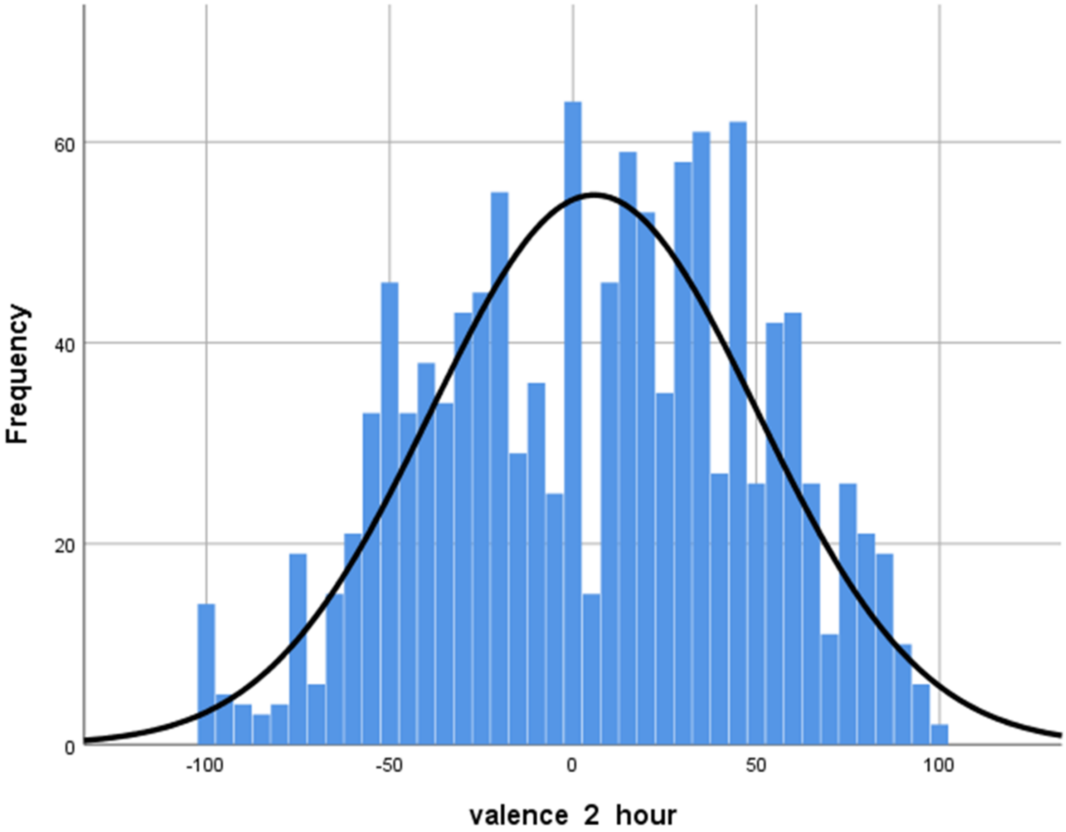
	Shapiro-Wilk		
	Statistic	df	Sig.
Mean SCL for 1-min interval	.067	504	.000
Arousal for 1-min interval	.945	504	.000
Valence for 1-min interval	.961	504	.000
Mean SCL for 2-hour interval	.136	504	.000
Arousal for 2-hour interval	.957	504	.000
Valence for 2-hour interval	.971	504	.000



Appendix B. Histogram with normal curve for SCL scores for the two-hour interval.



Appendix C. Histogram with normal curve for arousal scores for the two-hour interval.



Appendix D. Histogram with normal curve for valence scores for the two-hour interval.

Appendix E

Inter-individual correlations between SCL and arousal for the one-minute interval

			Mean SCL	Mean arousal
Spearman's rho	Mean SCL	Correlation Coefficient	1.000	-.114
		Sig. (2-tailed)	.	.596
		N	24	24
	Mean arousal	Correlation Coefficient	-.114	1.000
		Sig. (2-tailed)	.596	.
		N	24	24

Appendix F

Inter-individual correlations between SCL and valence for the one-minute interval

			Mean SCL	Mean valence
Spearman's rho	Mean SCL	Correlation Coefficient	1.000	-.217
		Sig. (2-tailed)	.	.307
		N	24	24
	Mean valence	Correlation Coefficient	-.217	1.000
		Sig. (2-tailed)	.307	.
		N	24	24

Appendix G

Inter-individual correlations between SCL and arousal for the two-hour interval

			Mean SCL	Mean arousal
Spearman's rho	Mean SCL	Correlation Coefficient	1.000	-.223
		Sig. (2-tailed)	.	.295
		N	24	24
	Mean arousal	Correlation Coefficient	-.223	1.000
		Sig. (2-tailed)	.295	.
		N	24	24

Appendix H

Inter-individual correlations between SCL and valence for the two-hour interval

			Mean SCL	Mean valence
Spearman's rho	Mean SCL	Correlation Coefficient	1.000	-.371
		Sig. (2-tailed)	.	.074
		N	24	24
	Mean valence	Correlation Coefficient	-.371	1.000
		Sig. (2-tailed)	.074	.
		N	24	24