Emotional valence as a mediating factor in emotional response coherence and measurement effectiveness of the Dual Process Framework.

Joël Hart
S1138413

University of Twente, GW faculty, Human Factors Engineering.

Supervisors:
Primary supervisor: Dr. Matthijs Noordzij
Secondary supervisor: MSc Erika van Wijk

UNIVERSITY OF TWENTE.

8-2-2016

Table of contents:

ABSTRACT. *This paper, building forth on of the work of Evers et al, attempts to find answers for the paradox of emotional response coherence. Emotions are often attributed as tools for the promotion of human survival, but replicating coherence of emotional responses in experiments often yields mixed results between different studies. This suggests that either the processing of emotional responses might be less consistent than previously anticipated, or that measurement irregularities are responsible for the incoherent results. Adopting a dual process framework methodology in analysis, separating the measurement of emotional coherence into two different systems of measuring, is speculated to improve the coherence of results. However, the correlation results found in the dual process framework research of Evers et al. seems to suggest that only moderate coherence connections exist and not all components of such a system seem to coerce evenly. Thus, the goal of this paper is to identify if another variable besides the method of data measuring and analysis influences response coherence. The particular variable of interest; Valence, is examined to possibly explain the variance in response coherence. Emotional valence is the specific direction or classification of emotion (negative vs positive types of emotion) and could be responsible for coherence differences as it determines the individual's likability of a stimulus. The second half of the paper covers an analysis testing if the dual process framework holds merit when applied to a data-set involving more participants and additional variables. In this data analysis it is also examined if the direction of emotional valence that participants encounter during the experiment can contribute to fluctuations in coherence. It was found that the basics of the dual process framework did not translate well to the used data set. Within system coherence calculations showed to be low and insignificant and only one situation specific effect of valence was supposedly found on between system coherence.*

Introduction

It's a well-established belief that beneficial emotional responses to specific situational contexts have aided our species in both its survival and growth (Darwin, 1872; Keltner & Gross, 1999; Neuberg, 2011). However, even though popular proposals like the Fight, Flight or Freeze theory claim that emotional responses as a whole should be of a consistent nature to promote personal survival (Cannon 1929; Ekman, 1992; Levenson, 1994), experimental research fails to produce uniform and confirming results when it comes to finding such coherence (Bonanno & Keltner, 2004; Reisenzein, 2000).

For a long time it has been assumed that emotional stimuli caused equal and consistent response fluctuations on all levels of emotional processing. For example, the mental experience of an emotion in a particular situation goes together with corresponding, stable changes in physiological and behavioural responses. However, research in this area currently seems to be moving away from this postulate and implies that either our situational responses to arousing stimuli do not cohere as consistently as previously assumed, or that the methods of measurement and analysis are incorrect.

As a reaction to the classical view of emotional response coherence, Evers et al. (2014) introduce a theoretical framework to deal with this paradox. According to their study, emotional responses can be categorized into two systems, the reflective and automatic. While these systems both relate to the different levels of emotional processing, they are relatively different from each other in how they measure emotional responses and their matching coherence. A more detailed elaboration on the Evers framework will be given later on in the paper.

Relevant to the previous statements on emotional response coherence, the theoretical segment of this paper sets two respective goals to be researched. First; I will reflect upon what constitutes to emotion and how traditional measurement methods succeed in accurately measuring it. Following this, Evers' dual process framework methodology is examined to see what the exact methodological difference are in this modern method and why they assumedly generates more reliable and valid data of coherence measuring. The Ever's methodology of using two separate systems to measure coherence is also tested in this paper. This is done by using the data gathered during another study using the Sing a Song Stress Test (SSST) (Brouwer, 2014; Derikx, 2015) and testing if measuring the response coherence of this dataset, containing far more respondents, is possible for both within system coherence and between system coherence.

In previous studies on emotions, variables have been found that are speculated to influence response coherence. The mediating and moderating variables in eliciting emotion that possibly affect response coherence are briefly examined by consulting existing literature. One of these will be of particular interest; this is the variable of perceived emotional direction, or “Valence”. Emotional valence is explored in depth during data analysis to establish if it significantly affects coherence scores. This is done by altering the existing data set in such a way that assumed positive and negative valence participants of the research are grouped separately on certain criteria variables containing a direction of emotion and by calculating the coherence scores of these groups separately to identify differences between them. The assumed discrepancy between the groups and the base coherence scores of the entire population of participants should indicate whether valence could influence emotional response coherence. The main goal of this paper is to establish whether the variable of valence influences the coherence of both the separate levels and between the different levels of emotional processing.

Theoretical background

What is emotion?

To start off, it is imperative to specify the definition of emotion used in this paper. Unfortunately, several hundred different definitions of emotion exist (Kleinginna & Kleinginna 1981), Therefore, to avoid further confusion I've chosen to adopt the description used in the glossary of the American Psychological Association (the APA). It refers to emotion as; “A complex pattern of changes, including physiological arousal, feelings, cognitive processes and behavioural reactions, made in response to a situation perceived to be personally significant” (APA, 2012).

Emotional response coherence

Emotional response coherence is a difficult concept to accurately describe. The idea of response coherence postulates that emotional reactions are consistent over the three levels of emotional processing; behavioural, experiential and physiological. (Mauss, Wilhelm & Gross, 2005) The behavioural level pertains to an individual’s behavioural reaction to a situation that elicits emotional arousal. For example; some specific facial expressions will follow a desired or unwanted stimulus. The experiential level involves the subjective cognitive and emotional experiences that someone encounters during arousing situations. Finally, there is the physiological level, which revolves around regulatory, physical changes in the body during an event that triggers an emotional response (Frijda, et al, 1992; Lazarus, 1991). In coherence research there is often referred to “within system coherence” and “between system coherence”. What is meant by these terms is the coherence of data on either a single separate level of (emotional) processing or between two or more levels. For example; if we only compare different experiential data measuring methods with each other in a coherence research, we speak of within system coherence. When the coherence of results on at least two or all three levels of processing is examined, we refer to it as between systems coherence.

Summarised; emotional coherence can be described as a pattern of reacting consistently to situations on the three different levels of emotional processing. If the measured span between an individual’s neutral and aroused state in a specific context does not tend to diverge much over or within different systems (Example: physiology vs. experience), we speak of emotional response coherence.

Measuring constructs of emotion

The most common methods of data measurement for constructs like emotions, stressors and the coherence between such variables can be grouped three categories (Scherer, 2005). These three consist of; using physiological measurements, accumulating experiential reports by respondents and conducting behavioural observation.

First off; by the application of an appropriate combination of measuring instruments it becomes possible to accurately quantify the physiological fluctuations an individual's body experiences as a result of a triggered arousing event. Changes in heartbeat rate and variability, skin conductivity, neurophysiological measures and blood pressure can relatively easily and reliably be monitored (Chanel, et al 2005). The second and perhaps most frequently used category of emotional response measuring is the use of self-report questionnaires to assess the individual's subjective feelings, experiences and perception. The individual can rate perceived emotions by using the provided scales, or by describing emotions verbally whenever appropriate (Caicedo & Beuzekom, 2005). And finally, there are observational behavioural measurements by using motor expressions. Through the measurement of, for example; facial expressions, gestures and alterations and fluctuations in voice tones (Dellaert, 1996) it becomes a feasible option to measure emotional responses of participants.

These categories for data measurements do all have their corresponding contextual advantages and disadvantages, (Harley, 2015) but are considered to be efficient and have been thoroughly researched. And although some instruments seem to be contextually more valid and reliable than others (Centre for studies on human stress, 2007), it is almost impossible to deny their fundamental usability. This creates the suspicion that not the measuring tools are to blame for incoherence, but perhaps the way data is processed in the analysis phase, which will be the main subject of research in the next section.

The dual process frameworks

The question remains how coherence experiments performed with reliable measuring instruments can produce inconsistent results. Especially if we regard the belief that it is necessary for emotional responses to be stable and coherent in order to effectively preserve a species.

A recent development in response coherence research came from Evers and colleagues (Evers et al., 2014). Evers acknowledges that response coherence research tends to be inconsistent and she makes the notion that a new method of measuring and processing data can lead to more desirable results. To achieve this, "the dual process framework" was introduced. This method divides emotional processing in two systems; a reflective system, which deals with the conscious experiences and behaviour of emotion and is generally measured by self-reported data. The second one is the automatic system, which includes the unconscious aspects of emotion, like physiological changes to arousing stimuli.

Evers theorized that both the systems would show significant correlation connections of within system coherence separate from each other, but would fail to show coherence when the results of the reflective and automatic systems were paired. This implies that emotions do not require corresponding changes between all three levels of processing in order to exist. For example, you can feel strong experienced cognitive reactions following a personal event, but this doesn't necessarily have to cause any physiological changes. The implications of this research show us that grouping all measured results of the three processing levels under a single variable might actually be the cause for incoherent results, as some emotional mechanisms can retain their baseline values while others do change heavily, depending on the stimulus.

To confirm their hypothesis, both the within and between coherence of the automatic and reflective systems were tested by administering two stress eliciting tests for each system to invoke emotions of anger on a sample size of 36 individuals. The data for the reflective system was gathered by self-questionnaires and data for the automatic system by using blood pressure and reaction time latency monitoring instruments, leaving out research and instruments regarding the behavioural component of emotional processing (Evers, 2014).

The results of Evers et al showed a statistically significant presence of within system coherence. The analysis yielded correlation scores of between .3 and .4, implying somewhat of a moderately strong connection. As also speculated by Evers, the source of incoherence in this area of research came from “faulty” data analysis regarding the grouping of multiple between system components (for example; the results of different measuring methods and instruments) at the same time. Between systems correlations failed to show a strong enough connection to prove that coherence exists between multiple systems. So, the main conclusion from this research is that not all levels of emotional processing have to change when an individual is exposed to a stressor or stimulus. It is noted by Evers et al. that not all related components of a specific system contribute to coherence equally well, which might explain the rather moderate connection in within system coherence. (Evers 2014).

It is important to highlight some specific aspects of the Evers research; The dual process framework divides emotional processing conform what was previously discussed in the current paper regarding the three processing levels of emotion. However, as mentioned before, the methods of observing emotional behaviour following an arousing stimulus were not included in the research and therefore the focus lies solely on physiological and experiential data measurements. This might impair the validity of between system coherence measurements as not all processing levels were included. Furthermore, the experiments of Evers showed that, while the within system correlations were significant, only moderate connections were found. It is also notable that a small sample trial was used. This implies that there surely is some merit to the theory, but research in this particular field isn't explored enough to exclude the possibility of other variables being responsible for these results. As Evers only specifically employed anger constructs, it would be a valuable endeavour to investigate if certain aspects of the chosen stimulus type can influence coherence (for example: If the stimulus is regarded as positive or not by the participants).

Later on in this paper, the dual processing framework methodology is tested by applying the reflective and automatic system methodology on a data set which also limits its research to physiological and self-reported data measurement, making it seemingly suitable for replication. However, this new dataset contains results gathered from a fivefold increase of participants, which creates the perfect opportunity to establish if larger sample sizes yield more significant coherence correlations.

Variables influencing the emergence of emotion

It appears to be unclear to what extent measurement and data analysis methods influence the formation of response coherence for both within and between system measurements. Because the present study is interested in the specific coherence of emotional responses, it seems logical to advance it in the direction of examining those components or variables that influence the process of eliciting emotional responses altogether. During research, a collection of interacting constructs involving emotional activation have been identified with the purpose of giving a basic overview of how emotional responses come to be. The constructs’ interrelation has been modelled and their role in the occurrence of emotion is briefly explained in the segment below. Figure 1 displays all the speculated variables that to some extent mediate and moderate emotional responses.

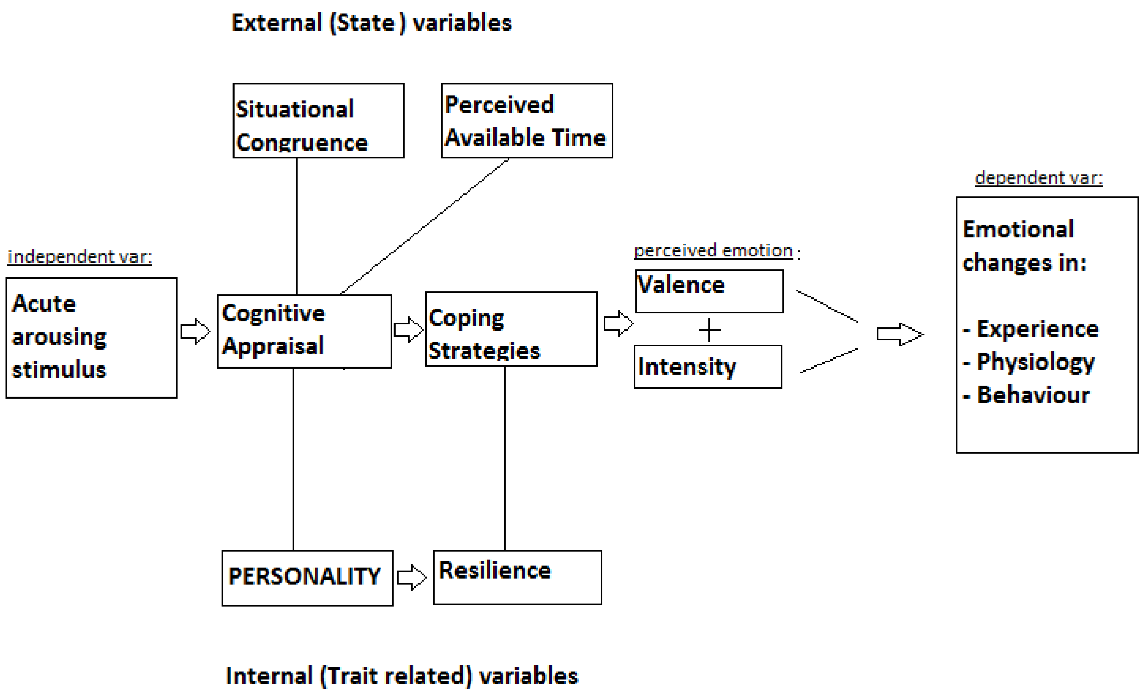

Figure 1: model of emotional activation

A distinction is made between “trait” and “state” related variables in figure 1. Trait constructs refer to stable aspects of an individual that are expected to remain the same. State variables are temporal and often mood dictated. Trait variables give an expectation on how an individual is likely to act in a given context, while state attributes often revolve around the influence of situation, environment or context on the individual’s appraisal, behaviour and feelings (Foundations of Sport and Exercise psychology, Weinberg & Gould, 2014).

An emotional reaction cannot exist without the appropriate relevant stimuli (In emotion research this is often, but not exclusively, a stimulus of a stressful nature referred to as “stressor”). Someone’s unaroused condition is referred to as its neutral base. Several factors influence how an arousing stimulus affects the individual, as described in the following section.

The model displays the trait related variable of personality. Personality encompasses several genetic and learned dispositions (Omel & Jeronimus, 2013). It can also be regarded as the long term counterpart of the more sporadic mood affected emotional state of an individual. Personality is a long lasting, coherent and internally stable “collection of behavioural dispositions” of an individual, which are unlikely to change. Personality also serves as a reliable pattern of how someone is expected to act emotionally. (Revelle & Scherer). Personality often determines a great portion of the emotional capacities an individual has for dealing with encountered stressors or other mentally taxing situations. This is also known as emotional resilience, a regulatory coping mechanism (Southwick, 2005).

A variable belonging to the state spectrum is situational congruence, which is the estimation whether or not a situation is emotionally relevant to the individual at all (Roidl, 2013). For example, the exact same situation can be interpreted differently at varying moments due to external influences of someone’s mood. Take for instance a person stuck in traffic. If this person has no immediate obligation or occupations, the situation will emotionally impact the individual less than the same situation occurring on a busy working day. (Roidl, 2013). This is also where the variable of available time comes into discussion. Especially in stressful situations constrained by approaching deadlines, more intense negative arousal is likely to occur in the form of anxiety and impair decision making and healthy functioning, causing emotional distress (Diederich, 1997).

Figure 1 implies that a combination between both the trait (personality) and state (congruence and time) variables, establish how a stimulus is appraised by the individual after perceiving the relevant stimulus. (Smith & Lazarus, 1990;) From the moment a stressor, or an otherwise arousing stimulus is perceived, coping capabilities and strategies are activated to mediate emotional fluctuations.

Nowadays, it is generally accepted that emotion consists of two directing variables; Valence and Intensity (Frijda, 1988). Valence refers to the individual's perceived direction of either attraction (positive valence) or aversion (negative valence) regarding a specific stimulus (Fossum & Barret, 2000). Valence is also considered to be a classification method used to catalogue different kinds of emotion. Intensity refers to the quantifiable levels of valence of a stimulus, so the emotional impact it has on the individual. Higher levels of intensity often invoke more extreme reactions to a stimulus in the appraised direction (Wintre & Vallence, 1993). For example, panic is an emotion with a strongly negative valence and high intensity. Serenity on the other hand has positive valence and low intensity.

When we assume that the variables describes in figure 1 operate the way as described, it is to be expected that in a completely identical experimental setting, context and situation, the same individual will react consistently and roughly achieve the same results when their mood is also similar. Thus we expect that replication of research will also show coherent replicated results. This should be the case when research is without any quantifiable differences besides operator/measurement errors. Due to their stability, it is highly unlikely that sustained trait variables like personality can be accountable for the incoherent findings in emotional response coherence on an individual level. If intra-individual differences are measured during (replicated) research, while the individual's trait variables remain unaltered, these incoherencies can only be attributed to either the influence of situational differences or the assumption that the three levels of emotional processing do not possess an all or nothing nature, as theorized in Evers (2014).

Emotional Valence and its influence on coherence according to literature

Valence describes the two possible different directions of emotional perception, negative and positive. Before I elaborate on the research goals of the valence-coherence data analysis study, a categorization model of both negative and positive emotions is given as a guideline for identification in figure 2. The fundamentals of this classification come from Ekman's theory of the six basic emotions, (referred to as BE) (Ekman, 1971; Ekman, 1999). Additional "sub emotions" (SE) were inferred from a multitude of sources (Xu, 2015; Plutchik, 2011) and classified under the corresponding basic emotional categories afterwards.

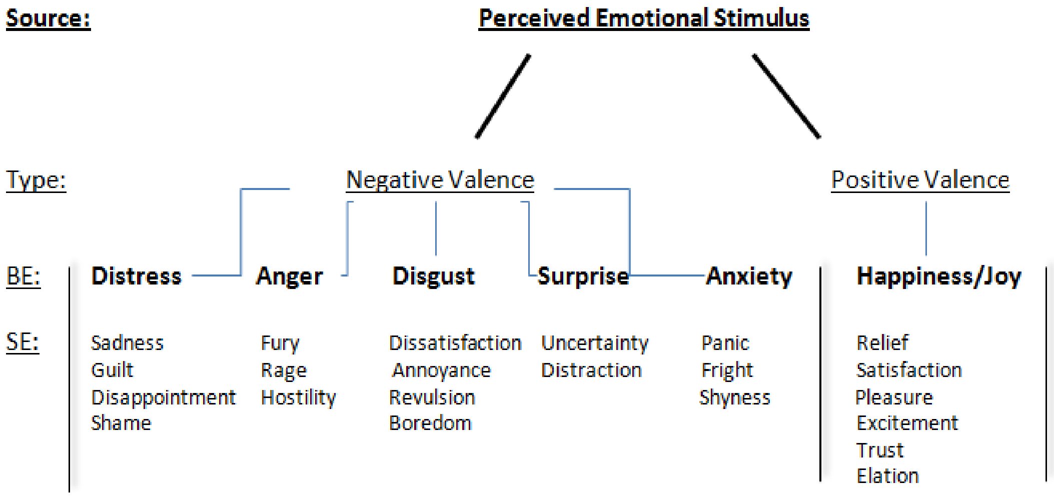

Figure 2: Valence classification

The reason why valence became a variable of interest for this paper were some of the studies conducted by Mauss and colleagues (Mauss, Wilhelm & Gross, 2005). Coherence of emotion was tested in this experiment with the emotion sadness, belonging in the distress category of negative valence. Results of this study implied that emotional responses elicited by sadness showed high association between the systems of experiential and behavioural processing, but only a very modestly between these systems and physiology. The authors additionally noted that they've previously found evidence that the negative valence types of surprise or anxiety tends to show less coherence between experiential and physiological responses than other emotional types do, because surprise and anxiety possess more substantial cognitive elements (Mauss, 2004; Reisenzein, 2000). These results imply that the valence of emotion can indeed have a diminishing effect on coherence if the particular categorized emotions contain more cognitive elements.

In accordance with the results of Mauss and colleagues, I assume that unexpected, cognitively taxing stimuli that possess a negative valence, do indeed produce less coherent results than positive counterparts. To test this, statistical data analysis will be performed by using the earlier mentioned data set. In the research that collected this data, participants were exposed to a stressful singing experimental stimulus. However, it is to be noted that a major limitation exists for analysis when using the current data set. The participants were not asked any questions that involved them to report which emotions they encountered during the measuring moments of the experiment. So the emotions that the respondents experienced are speculative, but based on reasonable expectations.

The nature of this study makes it plausible to assume that most respondents that are uninterested or unexperienced with public singing are more likely to encounter emotions of shame, surprise and anxiety during an experiment that "forces" them to sing in front of strangers. Such emotions do possess a more heavy cognitive load (Mauss, 2004), making it possible to test if response coherence is indeed lower for those participants with a negative valence. On the other hand, respondents were also asked if they do occasionally partake in singing as a hobby or in an organised community. It can once again be argued that most people who actively seek out the pursuit of singing would not mind to do so in an experiment and so maintain a more positive valence to the task. So a corresponding positive emotion of joy or perhaps constructs of pride can be attributed to the latter category of respondents.

To summarize, I expect that individuals who frequently engage in activities they enjoy, like singing, to maintain a more positive attitude when they have to unexpectedly perform these activities in front of strangers. Said individuals will be more likely to encounter positive emotions during comparable situations. This diminishes the cognitive load this group of individual’s experiences compared to those that do not seek out ventures like singing as a leisure activity. This results in the response coherence of the positive valence group to be higher than the negative valence group.

Hypothesis of the data analysis:

Data analysis should show if the assumptions regarding coherence that were introduced in the theoretical segment hold merit. This analysis must establish if comparable results to Evers' will be found on another data set, not original intended for coherence research, but with the inclusion of more participants and variables. For the general response coherence of the whole data set the following hypothesis is maintained:

> *“Significant coherence within the automatic response system will be found, while no significant coherence between the reflective and the automatic response system will be found.“*

The data set that is used in the upcoming analysis has another limitation preventing the current paper from delivering complete research: There is only one instruments of self-reported measurements included for usage in the dataset's study. This means that it is impossible to do a within system coherence analysis of the reflective system because there is no way to compare different instruments. Unfortunately, because of this the Evers methodology cannot be replicated entirely and just it’s basic principles of system specific analysis are maintained. Within system coherence of the physiological automatic system can still be determined, as multiple measurement methods exist for this category in the research. Between systems coherence can also be analysed by correlating data of one of the instruments measuring physiological state on different moments, with self-reported data corresponding to the same measuring moments.

Once the attempt to replicate the Ever's method is complete, and the coherence scores mentioned above have been established, a general overview of coherence for the whole group should be produced. Next, participants will be divided in separate groups based on estimations of emotional valence. One group is speculated to have a negative emotional valence, because they experience negative emotions due to the experiment's singing condition and the other group is assumed to have a positive valence, due to more familiarity with the experimental condition. The hypothesis for the research goal of establishing if valence alters coherence is:

> *“The negative valence group will show less physiological within-system and between-systems coherence than the positive valence group.”*

The expectation is that comparable coherence results as in Evers' study will be found and that valence does indeed noticeably influence such coherence scores.

Methods

Participants

The data set used for the data analysis came from a study regarding the existence of personality profile types in stress (Derikx, 2015). For this set of data, 154 participants were enlisted. The following segment will reflect on the methods of this study and how the analysis of this paper was set up.

To recruit the participants, an online study recruitment instrument was chosen to be the sole tool of enlisting subjects. The utilised program was the Sona Systems application with which the participants themselves voluntarily sign-up. The deployed version of the web application was specifically licensed for use by University of Twente students and other members involved with the university. Compensatory study credits were granted to students who completed the experiment. Of the 102 individuals recruited, the gender ratio was 39 males versus 63 females. A significant portion of the participant pool appeared to consist of students, evident from the age range which spanned from 18 to 55 years old (with a mean = M = 22.13 and standard deviation = SD = 4.393). The majority of participants had the Dutch nationality, with the exception of 35 of the participants being German. Two participants were excluded from the results post study, due to possessing ailments that might bias or distort results.

During the study, participants were initially informed that they were taking part in a study measuring data of personality and corresponding levels of fitness. In truth, this was a deception fabricated to guarantee that the measurement of the actual variable of interest could unobtrusively transpire without the interference of bias or knowledge effects from preparation by the participants. The nature of the sing along stress test (From now on referred to as SSST) requires the participants to be exposed to an unexpected stressful situation for the data to hold significance. Fitness level was chosen to be the deceptive substitute variable because this would believably explains the use of heartbeat and skin conductance measurement instruments during the study, which are in truth necessary to obtain data on the actual variable of stress.

As a final disclaimer: The experimental conditions and all operations performed in this study were approved by the ethics board committee of the University of Twente as long as ethical conditions were met. Naturally, participants were to retain complete independent freedom during the study and had no forced obligations to the experimenters. Meaning that it was made explicitly known to the participating individuals that they were fully permitted to terminate participation or refuse cooperation at any given time. Additionally, no expedient or intentional harm should befall a participant due to the actions of experimenters. All these aspects were specifically documented in the consent form presented to and required to be signed by participants prior to the study.

Apparatus and materials

To administer tests, a Windows computer desktop attached to a flat screen monitor was used to convey information. The computer equipment was controlled by a standard mouse and keyboard method. The distance between the monitor and participant was approximately 60 cm diagonally, enabling a healthy straight back position for the participant. Programming language software from the Python Software Foundation was used to create a time indicator when connected to voltage isolator as well as displaying information regarding the assignments and simultaneously recording a participants subjective stress reports. Said voltage isolators enabled accurate starting times and timeframe stamps regarding the assignment.

The physiological measurements were performed by equipment from Biograph Infiniti. The section below will highlight and elaborate all the provided equipment and accordingly for what type of physiological measurement it was utilised.

The measurement of a participant's skin conductance was conducted by two finger sensory measurement equipment from the SCR (skin conductance response) Biograph Infiniti package. They are attached by Velcro wrap on the ring and index finger of the left hand. The sensors consist of silver ionized buttons (Ag-AgCl). Din to snap adapters connect the sensors to the G-tec amplifier. The recording frequency was 256Hz sensor supply voltage was 7.3 mV.

The EDA wrist sensor, not part of the Biograph Infiniti package, was synchronised with a computer to be able to later match the starting time with the assignment windows. The wrist equipment's segments are not marked by the voltage isolator pulses. Segments are still effectively identified based on the start time, logged on the same computer the wrist sensor is synchronized with both the skin conductance response and heartrate sensors and converge in two cables that are inserted in the amplifier in port C (SCR) and port E (HR). The voltage isolator is inserted in port (H). Sensors attach to the ProComp Infiniti 8 channel amplifier in ports measuring at 256Hz (C through G) and provide anti-aliasing filter (5th order Butterworth, 30dB typical rejection). The amplifier is connected to the identical laptop used to administer the personality tests via the TT-USB-T7700 that transforms the fibre optic signal to USB.

Heartrate monitoring equipment, the EKG package, was also supplied by the Biograph Infiniti. This package contains three removable silver ionized (Ag-AgCl) button sensors that make contact with the skin. The DIN to SNAP adapter wires leading to the amplifier are colour coded (yellow, blue and black) and need to be attached to the participant using re-usable medical grade non-latex tourniquets. The black sensor is a ground and is attached in such a manner that the sensor makes contact with the radiant side of the left arm. The blue sensor is attached to the same tourniquet and makes contact with the centre of the left wrist. The yellow sensor has its own tourniquet and is fastened to make contact with the centre of the right wrist. Recording frequency was 256 Hz and sensor supply voltage was 7.3mV

Multiple tests were used in this experiment; The NEO-FFI and IRS tests were provided to the experiment by TNO via web browser access. The NEO-FFI consists of 60 items measuring the 5 spectrums of personality. The IRS test measures data on general stress and coping ability. All test utilize 5 point Likert scales with the scoring range of 1 = Completely Agree to 5 = completely disagree. The results from participants were processed by the TNO and returned to the researches after analysis.

The online survey and questionnaire tool website "thesistools" was used to create the demographic survey. This questionnaire included elements regarding lifestyle choices and health related variables of the participant. The participants utilized the google chrome browser to complete all surveys and questionnaires. Due to the majority of the material being in Dutch and the deviating German participant population not being substantial enough, a mediating solution for the language barrier was presented in the form of a wordlist, translating especially difficult words. Thesaurus websites were also allowed access to during the experiment.

Procedure:

After the Sona Systems signup trial had been completed and a physical appointment was made, the participants were individually instructed to be present at the designated location on the Utwente terrain at the appropriate time. The experiment was to be conducted in a specified room in the research laboratory. After the participant was present on location, an introductory briefing was held informing the subject of the (deceptive) goal of the study; testing personality factors in relation to fitness levels. This is also the stage were, for the sake of anonymity, the participant received an unique reference number.

Before experimentation could commence, first the ethical requisites had to be met. To this end, the participant was asked to thoroughly read the informed consent form in which the possible risks and hazards were stated that could occur while participating. The consent form also included standard practice information. Seeing as the nature of the sensory equipment demanded some physical experimenter-participant contact to be effective, a “personal contact form” was also presented to the participant, informing him or her of the necessity and severity of this contact. Presenting these forms is a standard, but mandatory procedure for all research conducted at the university and is a requisition created by the ethics committee board of the Utwente. Both forms needed to be considered and signed by a participant before any experimental research could commence.

After consent was given the participants were invited to fill out two questionnaires by using a laptop. One questionnaire revolved around demographical data and was mainly used to identify those factors that might threaten validity. The second questionnaire consisted of two tests; the IRS and NEO-FFI tests. These tests were provided to the research by TNO. The estimated time of completing both these questionnaires was no more than 30 minutes. After receiving global instructions on how to fill out the forms and, if appropriate giving German participants an additional wordlist to translate the Dutch terms, the participant was left unattended by the experimenter to independently complete the questionnaires. The experimenter would still be present in the adjacent room keeping watch over the participant by a closed camera system. The participants, however, were instructed that they could bring in the experimenter for aid or questions by knocking on the door. This was also the procedure to let the experimenter know that they completed this section of the research.

After wrapping up the questionnaires, the research leader instigates the experimental condition of the research by bringing in an (unknown to the participant) accomplice of the experimenter acting as a second participant. The condition is that this second participant came in early and is therefore also seated in the same room as the actual participant. To reinforce the ploy, the accomplice is given similar instructions as the real participant and is also asked to fill out questionnaires. At the same time the participant is moved to the main computer and is in turn linked to the sensory equipment. As the EDA gauge requires time to effectively make contact, this part of equipment is attached first. This gauge is akin to a substantial wristwatch. Heart rate monitors and the SCR sensors are subsequently attached to the participant. It is essential that a person's left hand and wrist are completely devoid of additional accessories to attach the physiological measuring equipment.

The participant is advised to remain seated in a comfortable sitting position as movement is essentially disruptive during the following 10 minutes. Meanwhile, the researcher initiates the Python software displaying the instructions for the participant while the accomplice starts recording the experiment as it commences. The experiment gives additional faux instructions to the accomplice which remains seated near the real participant. The experimenter remains in the room to ensure that no excessive movement is conducted by the participant.

The participant is subjected to multiple stages in which he or she has to prepare or sing a song according to the SSST. During this time the experimenter has to be extra vigilant to prevent the participant from moving. Especially during the singing stage some corrections and encouragement posed to be essential. To invoke a valid reaction by the participant, the accomplice was required to keep up his or her role. Tasks of the accomplice include acting accordingly to stimuli during the experiment, pretend to fill out the demographic and test forms like the participant did moments earlier and starting the recording unobtrusively.

Baseline and subjective stress measurements are recorded automatically post-test completion and the accomplice registers whether the participant actually partook in singing a song. The experimenter removes the technical sensory equipment from the participants and expresses their gratitude. Data is stored according to the corresponding identification number of the participant. A debriefing takes place in which the participant is informed of the applied deception and what the true purpose of the experiment is in relation to the SSST. The participant receives the option to leave their contact information to stay frequently updated about the research and it's outcome.

<u>Data analysis procedure:</u>

For statistics data analysis procedures, IBM SPSS statistics data editor version 21 was used. The research data of the 154 participants was categorized into appropriate variables and coded accordingly so that statistical analysis of the data became feasible.

In the data analysis, two major coherence tests will be performed. The first includes establishing both within system and between system response coherence scores for the whole participant pool. Between system coherence is calculated by correlating variables of the reflective system with the automatic system. This is done by performing bivariate correlations of self-reported change scores with the corresponding physiological change score data belonging to a certain measuring moment in the experiment. Because only a single self-reported instrument was used in this research to measure experiential reactions on specific measuring moments, it becomes unable for this research to perform data analysis on within system scores for the reflective system, as at least two different instruments are required for this. For the automatic system, this can be analysed by taking data of skin conductance peaks change scores and correlating these with heartrate data belonging to the same measuring moment.

First off, to realize this analysis, change scores are calculated from the raw data. The raw data contain information on corresponding times during the experiment. To determine change scores, the difference between one measurement moment and the next measurement moment was calculated. This is done for heart rate, skin conductance peaks and self-reported stress data (This is performed by using the calculations shown in Appendix B). There are five measuring phases in total that utilize both physiological and self-reported instruments. These phases are identified as: “Neutrale zinnen”, “Voorbereiding zingen”, “Zingen” and two “Post baseline” measurements.

The raw data of all phases contain a single measurement of one self-reported instrument excluding the pre_baseline fase, which has no self-reported measurement. Four measurements of both heartrate and skin conductance peaks data are present in both the first two phases (pre_baseline and neutral) and two measurements in each of the last phases (post_baseline). The preparation and singing phases both contain a single measurement of each instrument for these variables. These physiological measurements assess the same data over the span of the two minute, thus each measurement contains an equal amount of 30 seconds of data (the post baseline phases last 1 minute each, this is why these phases only include two physiological measurements). Because physiological changes have likely stabilized the most in every last physiological measurement, only these last measurements (either referred to as #4 for most phases and #2 in the post baseline phase 1) should be used when processing raw scores. The ‘preparation for singing’ and the ‘singing’ phase both last 30 seconds, so each physiological variable has one 30 second measuring interval in both these phases. Physiological instruments are used during an additional phase measuring pre_baseline data.

The calculated change scores between phases are referred to as: SCR_change_0, SCR_change_1, SCR_change_2, SCR_change_3, and SCR_change_4 for skin conductance peaks. To illustrate an example; the first change score category (SCR_change_0) belongs to the differences between pre baseline physiological data and those of the “neutrale zinnen” phase. This goes on until the final difference category (SCR_change_5); the one between the end of the “zingen” phase and the end of post baseline data. For heart rate data, the same principle applies; the change score variables are called: HR_change_0, HR_change_1, HR_change_2, HR_change_3, and HR_change_4. For the pre-baseline phase there are no self-reported measures, so no differences exist for this phase. This lead to self-reported data only having 4 categories (pre baseline is excluded); Selfreport_change_1, Selfreport_change_2, Selfreport_change_3 and Selfreport_change_4. This means that the first change score of self-reported data is between the phase “neutral zinnen” and “voorbereiding zingen”.

For between system coherence it was chosen to compare self-reported (first mentioned variable) data with data of heartrates (second mentioned variable): To accomplish this, we use change scores of self-reported stress (Selfreport_change) and change scores of heart rate (HR_change). These variables are correlated to show coherence scores. The heartrate variable is interchangeable with skin conductance data of the corresponding moment (for example HR_change_1 should be switched with SCR_change_1 etc. to measure the coherence of the different physiological instrument).

The data belonging to the variables that are to be used for change scores coherence calculation are subjected to normality tests. This is done to establish if correlational analysis is the appropriate statistics method. The Shapiro-Wilk test is performed over the variables to establish if data is normally distributed. These tests are performed for heart rate, skin conductance peaks and self-reported stress.

For the data analysis about the influence of valence on coherence, first the participants must be divided into two separate groups. This division is based on criteria whether or not the participant is experienced and familiar enough with (public) singing. The respondents experienced with singing are put in group 1 and the remaining ineligible participants are grouped together in group 2.

The criteria used to group individuals are the responses given by all participants on five different inquiries that measure singing experience in the research's questionnaire. These five variables are: "Singing", "Singing_Presence_Others", "Singing_experience_Public", "Singing_experience_Solo_Public", "Singing_experience_Acapella". The variable of "Singing" inquires the participant how often they engage in any singing related activity. The participant had the option to select one out of five possible replies measuring singing activity ranging from scores: 1 = "Never" to 5 = "Often". The remaining variables respectively determine if participants have experience with; singing in the presence of others, singing in public with the support of additional singers, singing in public alone and singing in an acapella style setup. The participants are able to either answer "yes" (= option code 1) or "no" (= option code 2) on these questions.

Prior to the analysis, a few cut off criteria were determined to decide in which cases a participant was deemed experienced enough with public singing to be administered in group 1. Participants that answer the first question; "Singing" with option 1 - "Never" are automatically administered to group 2 because it can be assumed they have no experience with singing at all and thus do not possess a positive valence for public singing. 19 participants of the dataset were placed in group 2 based on this criterion, plus the supporting find that they also all specified to have no experience with public singing whatsoever by answering the other relevant questions with "no".

For the rest of the participant pool: Every remaining participant does have to have some experience with public singing in order to be grouped in group 1 and has to specify this by answering "yes" on at least one of the remaining experience measuring questions. Maintaining this criteria, another 47 participants that specified that they "rarely" (= option 2), or "on occasion" (= option 3) engage in singing activities on the "Singing" question and also implied that they have no experience with public singing (measured by them answering only "no" on all the other question), were put in group 2.

Of all the participants that answered the question "Singing" with either "regularly" (=option 4) or "often", only one participant reported that they do not have any experience with public singing. Because a notion of experience with singing in public is required to suggest the ability to maintain a positive valence during the experimental condition of the research, this individual was grouped in group 2. This is because it is uncertain if this individual does indeed possess a positive attitude for public singing, regardless of how much they seem to engage in singing. Generally it seems that those most frequently involved in singing activities do tend to seek out at least one source of public singing. However, when this is not the case, we advise to place those participants in the unexperienced group. This makes the total of participants included in the somewhat-to-heavily experienced group to

87. The remaining 67 participants are placed in group 2 because they have no discernible experience to speak of regarding singing. The two groups both get their own data file for further analysis.

After the separation of participants has been completed, the two individual group data files are exposed to the exact same statistical analysis procedures that were done for response coherence correlations of the whole participant pool. During analysis procedures, some participants are automatically excluded during the research by the software because their data is incomplete and thus ineligible statistical analysis.

Results

The separate data files of the three groups are each subjected to a normality test and three bivariate correlation analyses to answer the research goals. The results are listed and described in this section.

Data set 1: whole group analysis results.

The normality test was performed in the form of the Shapiro-Wilk test, due to the relatively small size of the sample ($n < 2000$). The lowest W found was .820 for HR_change_4 ($p < .001$). The highest W found was .966 for HR_change_3 ($p < .01$). All results were significant at least at the $p < .01$ level. This means none of the tested change scores were perfectly normally distributed. However, they were found to sufficiently approach the normal distribution to proceed with the correlational analyses. The physiological variables that most approached the normal distribution were the change scores 2 and 3, which correspond respectively to the change score between preparing to sing and singing, and the change score between singing and right after singing ($.96 < W < .97$, $p < .01$).

Using correlational analysis comparing HR_change and SCR_change , the significance of found correlations varied between $p = .524$ and $p = .965$. The correlations varied between $r = -.004$ and $r = .057$. Consequently, no correlation can be assumed to exist between heart rate change and skin conductance change in this study.

Some significant correlations were found between heart rate change and self-report change. A correlation of $r = .227$ ($p = .012$) was found between HR_change_1 and Selfreport_change_1, a correlation of $r = .243$ ($p = .007$) was found between HR_change_3 and Selfreport_change_3, and lastly, a correlation of $r = .187$ ($p = .038$) was found between HR_change_4 and Selfreport_change_4. Only the correlation found in timeframe 2 was not significant. Apparently, three out of four timeframes contain a weak correlation between heart rate and self-reported stress.

Only one significant correlation was found when comparing SCR_change and Selfreport_change. This correlation was $r = .173$ ($p = .047$), and was found between SCR_change_3 and Selfreport_change_3. Therefore it can be assumed a weak correlation exists between skin conductance response and self-reported stress in timeframe 3.

Data set 2: Experienced singers

Again, the Shapiro-Wilk test was performed to test the normality of the distribution of data. The lowest W found was W = .831 for HR_change_0 ($p < .001$). The highest W found was W = .959 for SCR_change_2 ($p < .05$). All results were significant at least at the $p < .05$ level. This means none of the tested change scores were perfectly normally distributed. However, again, they were found to be sufficiently close to the normal distribution to proceed with the correlational analyses in favour over another method.

The results of correlational analysis comparing skin conductance change and heart rate change showed no significant results. The found significance levels varied from p = .383 to p = .969, while the corresponding correlations varied from r = -.107 to r = .100. Consequently, no correlation between heart rate and skin conductance can be assumed to exist in the group of participants with singing experience.

Heart rate change and self-report change were assessed next. Two significant correlations were found as shown in table 1. The first significant correlations was found at timeframe 1 and the second significant correlation was found at timeframe 3. These results suggest that HR_change and Selfreport_change are modestly correlated during timeframes 1 and 3.

Table 1. Outcome of between system coherence correlation analysis of heartrate and self-reported differences for the positive valence group over the phases of the research. Correlation values (r=) are displayed together with the corresponding significance (p=). Relevant values are those between corresponding timeframes (1-4).

Change score variables	**Heartrate changes 1**	**Heartrate changes 2**	**Heartrate changes 3**	**Heartrate changes 4**
Selfreport changes 1	*r*= ,278 *p*= ,026	*r*=-,030 *p*= ,810	r= -,008 p= ,951	r= -,087 p= ,490
Selfreport changes 2	*r*= -,135 *p*= ,333	*r*= ,098 *p*= ,434	r= -,114 p= ,361	r= ,129 p= ,305
Selfreport changes 3	*r*= -,123 *p*= ,026	*r*=-,356 *p*= ,003	r= ,393 p= ,001	r= -,086 p= , 497
Selfreport changes 4	*r*= -,127 *p*= ,317	*r*= 002 *p*=,988	r= ,108 p= ,389	r= -,118 p= , 349

During the analysis of the correlations between skin conductance change and self-report change, one significant correlation was found. This correlation was found between SCR_change_3 and Selfreport_change_3, and had the value r = .252 (p = .031). The correlation between

SCR_change_2 and Selfreport_change_2 was nearly significant, at $p = .058$ and with $r = -.223$. A weak correlation between skin conductance change and self-report change during timeframe 3 is assumed to exist. A correlation between these two variables during timeframe 2 is somewhat likely, but cannot be assumed due to insufficient significance.

Data set 3: Participants without singing experience

This dataset too was subjected to the Shapiro-Wilk test of normality. The lowest W found was $W = .735$, and the highest W found was $W = .983$. Three values of W were not statistically significant: HR_change_2 had $W = .969$ with $p = .153$, HR_change_3 had $W = .983$ with $p = .617$, and Selfreport_change_1 had $W = .962$ with $p = .078$. All other values of W were significant at the $p < .05$ level. All but one value of W were higher than .847. The three statistically insignificant variables mentioned above can be assumed to be normally distributed for participants without prior singing experience. Importantly, the other variables still approach the normal distribution sufficiently to proceed with correlational analysis.

In determining the correlations between skin conductance change and heart rate change, again, no significant correlations were found. The correlations that were calculated varied between $r = -.074$ and $r = .137$. Significance levels varied between $p = .316$ and $p = .873$. Apparently, no correlation exists between skin conductance change and heart rate change at any interval of this study with participants without prior singing experience. As no other significant correlations are found on this subject across the three groups, this concludes a pattern of the absence of within system response coherence

One significant correlation was found between heart rate change and self-report change. Between HR_change_4 and Selfreport_change_4 a correlation of $r = .271$ was found ($p = .043$). This means there likely exists a correlation between heart rate change and self-report change during the last timeframe of the experiment, in the group of participants without prior singing experience.

Lastly, not a single significant correlation was found when analysing skin conductance change with self-report change. The levels of significance that were found varied between $p = .164$ and $p = .743$, and the correlations varied between $r = .044$ and $r = .185$. Consequently, no significant correlation can be assumed to exist between skin conductance and self-reported stress in participants with no prior singing experience.

Discussion

The current paper initially stated two research goals: To see whether it was possible to calculate coherence in the available data set by dividing the analysis of data into two systems; The reflective and the automatic, as introduced in the Evers dual process framework method. The second goal is to determine if coherence scores are influenced by the type of valence, or the positive or negative perceived emotions by individuals following a stimulus.

All the results found regarding within system and between system coherence contradicted the findings of Evers' research. Results of data analysis of coherence within the automatic system did not show any significant correlations between heart rate and skin conductance response. This means that no conclusive statements can be made regarding if the (weak) correlations that were found are valid and reliable. This is rather surprising, as at least some significant results of within system coherence were expected. A study done by Evers et al. (2014) did find such a coherence.

With the insignificant results that the current research produced, it is not possible to give any conclusive verdict about the first research question. Because of the lack of confirming results, it is to be concluded that we can't answer whether the basic methods of the dual process framework can or cannot be applied to a data set that utilises the same type of measuring methods based on the current research. Due to these insignificant findings, it also remains unclear if a data set with more respondents and additional variables could potentially produce more significant results of within system coherence. Furthermore, based on the results we cannot prove or disprove that the between system coherence of such a measure should yield low connections as theorized by Evers, once again due to the insignificance of results. If a weak correlation does in fact, exist between HR and SCR, it is possible the current sample size was simply too small to significantly show its existence.

Another unexpected finding is that between system coherence was not always found to be low and insignificant for this data set. For the whole group and the positive valence group, the change score correlations of both the physiological instruments with the self-reported instrument data, showed significant connections on the third measuring moment (which includes a measurement span of the participant singing until just after the singing had been completed). This implies that a coherence connection between self-reported stress and physiological measurements exist between the phase of having to sing to the singing being over for these groups. This effect was found to be stronger for the positive valence group, and the effect was not found in the negative valence group. A correlation of about .4 was found between heart rate and self-reported data for the correlated change scores while this was just short of .25 for the whole group. Strangely, no such significant scores were found in this phase for the negative valence group. Assumedly, the correlation found in the total group was caused by the data from the positive valence group.

Still, the research question whether valence influences coherence is also hard to answer with the current results, as the dual process frameworks methods did not seem to be compatible with this data set for the general population. A few systematic findings of significant correlations were found for between system coherence, which could indicate a specific, situational connection of coherence. These findings, however, were more specific than anticipated. Only two (respective) measurement moments of the study showed some correlation: It seems coherence between systems is higher in the timeframe after participants finish singing, but only in participants with some positive valence towards singing in front of others. Clearly, positive valence had some influence on the current results. The demonstrated interaction between the stage of the experiment and singing experience seems like a promising lead to further exploring this line of research.

Limitations

Some limitations have to be mentioned that might have impaired this research from reaching its full potential. First off, replicating the Evers' dual process framework method was not entirely feasible as this study used only one reflective measurement instrument, and therefore did not allow the calculation of coherence within the reflective system. Additionally, participants were not asked about the type of emotion they experienced during the different phases of the experiment, so these had to be deduced from speculations partly stemming from other variables of the research (whether participants sang often and did so publically was linked with a more positive attitude towards the Sing a Song Stress Test). The two groups were separated based on deliberate assumptions. Furthermore, the criteria for how participants were to be grouped might not be an optimal way to reliably measure either negative or positive valence.

Future research:

As response coherence research is prone to inconsistencies (Evers 2014), more specific research utilizing the dual process framework could give further insights on how response coherence works and the usability of this method. The mentioned limitations make it impractical to accurately integrate the methods of Evers into the data analysis using this set of data. In future research, some of these hurdles could be eliminated to provide more effective research. Such research should able to produce more satisfying results. For continued research in the area of the valence/emotional response coherence relation, I would recommend to create a specific experiment in which it is more easy to identify the valence of participants. The Sing a Song Stress Test seemed suitable for valence tests, but it was rather hard to establish which participants truly had a positive attitude regarding public singing. It would also be advisable to add items to self-report questionnaires specifically asking the participant what kind of emotion they are experiencing during certain moments of measuring (perhaps this can be mediated by giving participants an exemplary list of emotions, as not everyone is familiar with the extent of emotional categorization). Doing so would allow a more smooth transition for researchers to create different participants groups based on valence. the inclusion of more different self-reported measuring tools to enable the analysis of reflective between system coherence.

Conclusion

Although the data analysis did not succeed in giving a clear conclusive answer on the research goal questions, some interesting connections regarding between system coherence and valence were found. It seems that between system emotional response coherence can exist over situations in which an individual has to perform a certain (somewhat desirable) activity, until the end of said activity. As this effect was found to be stronger correlated in a group presumed to possess a more positive valence than a base group containing all participants, this can indicate that valence can indeed improve coherence results under the right circumstances. More research devoid from limitations is needed to confirm these suspicions.

References:

Bonanno, G. A., & Keltner, D. (2004). The coherence of emotion systems: Comparing “on-line” measures of appraisal and facial expressions, and self-report. *Cognition and Emotion, 18*, 431-444.

Brouwer, A.M & Hogervorst, M.A. (2014). A new paradigm to induce mental stress: the Sing-a-Song Stress Test (SSST). *Frontiers in Neuroscience*, 8(224).

Caicedo, D.G & Beuzekom, van, M. (2006). “How do you feel?” An assessment of existing tools for the measurement of emotions and their application in consumer products research. *Delft University of Technology Department of Industrial Design.*

Candland, D. (2003). The persistents problems of emotions. Em D. Candland, J. Fell, E. Keen, A. Leshner, R. Plutchik, & R. Tarpy (Orgs.), *Emotions* (pp. 01-84). Nova York: Authors Choice.

Cannon, W.B. (1929). ORGANIZATION FOR PHYSIOLOGICAL HOMEOSTASIS . P*hysiological reviews* , 9(3) .

Centre for studies on human stress. (2007). *Physiological Measures*. Retrieved 13 January, 2016, from http://www.humanstress.ca/researchers/measure-stress/physiological-measures.html

Chanel, G., Kronegg, J., Grandjean, D., Pun, T. (2005). Emotion Assessment: Arousal Evaluation Using EEG’s and Peripheral Physiological Signals. University of Geneva: Computing Science Center.

Darwin, C. R. 1872. *The expression of the emotions in man and animals*. London: John Murray.

Dellaert, F, Polzin, T.S & Waibel, A. (1996) A. Recognizing emotions in speech. In ICSLP.

Derikx, L. 2015. *STRESS, COPING AND PERSONALITY.*

Diederich, A. (2003). Decision making under conflict: Decision time as a measure of conflict strength. *Psychonomic Bulletin & Review*, 10(1), 167-176.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), Nebraska Symposium on Motivation, 1971 (Vol. 19, pp. 207–283). Lincoln: University of Nebraska Press.

Ekman, P. (1992). An Argument for Basic Emotion. *cognition and emotion*, 6(3/4), 169-200.

Ekman, P. (1999). "Basic Emotions” In: T. Dalgleish and M. Power (Eds.). *Handbook of Cognition and Emotion*. John Wiley & Sons Ltd, Sussex, UK.

Evers et al. (2014). Emotion response coherence: A dual-process perspective. *Biological Psychology*, 98(43-49).

Fossum, T., & Barrett, L. F. (2000). Evaluation and description in the personality-emotion relationship. *Personality and Social Psychology Bulletin*, 26, 669-678.

Frijda, N.H. (1988). The Laws of Emotion. *American Psychologist*, 1(43), 349-354.

Frijda, N.H., Ortony, A., Sonnemans, J., & Clore, G. (1992). The complexity of intensity: issues concerning the structure of emotion intensity. In Clark (Ed.), Review of personality and social psychology (Vol. 13, pp. 60-89). Beverley Hills: Sage

Harley, J. M. (2015). Measuring emotions: A survey of cutting-edge methodologies used in computer-based learning environment research. In S. Tettegah & M. Gartmeier (Eds.). *Emotions, Technology, Design, and Learning* (pp. 89-114).London, UK: Academic Press, Elsevier.

Keltner, D & Gross, J.J. (1999). Functional Accounts of Emotions. *cognition and emotion*, 13(5), 467-480.

Kleinginna, P.R & Kleinginna, A.M. (1981). Motivation and Emotion. *A categorized list of motivation definitions, with a suggestion for a consensual definition*, 5(3), 263-291.

Lazarus, Richard S. (1991). Progress on a cognitive-motivational-relational theory of Emotion. American Psychologist, 46(8), 819-834.
Levenson, R.W. (1999). The Intrapersonal Functions of Emotions. *cognition and emotion*, 13(5), 481-504.

Liu, X & Xu, Y. (2015). Relations between affective music and speech: evidence from dynamics of affective piano performance and speech production. *Frontiers in Psychology*.

Mauss, I. B., Wilhelm, F. H., & Gross, J., J. (2004). Is there less to social anxiety than meets the eye? Emotion experience, expression and bodily responding. *Cognition and Emotion*, 18(5), 631–662.

Mauss, I. B, Levenson, R. W, Wilhelm, F. H, McCarter, L, & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, 5(2), 175–190.

McAdams, D. P. (2009). *The person: An introduction to the science of personality psychology.* Hoboken, NJ: John Wiley & Sons, inc.

Neuberg, S.L, Kenrick, D.T & Schaller, M. (2011). Human Threat Management Systems: Self-Protection and Disease Avoidance. *Neuroscience & Biobehavioral Reviews*, 35(4).

Ormel, J, Jeronimus, B.F, Kotov, M, Riese, H, Bos, E.H, Hankin, B. (2013). "Neuroticism and common mental disorders: Meaning and utility of a complex relationship". *Clinical Psychology Review* 33 (5): 686–697

Plutchik, R. "The Nature of Emotions". American Scientist. Retrieved 14 April 2011.

Reisenzein, R. (2000). The subjective experience of surprise. In H. Bless & J. P. Forgas (Eds), The message within; The role of subjective experience in social cognition and behavior (pp. 262-279). Philadelphia, PA: Psychology Press

Revelle, W & Scherer, K.R. (Unpublished). Personality and Emotion. *Oxford Companion to the Affective Sciences Oxford University Press*.

Roidl, E, Frehse, B, Oehl, M, Höger R. (2013). The emotional spectrum in traffic situations: Results of two online-studies. *Transportation research part F: traffic psychology and* behaviour 18, 168-188

Scherer, K.R. (2005). What are emotions? And how can they be measured?. *Social Science Information*, 44(4), .

Smith, CA. & Lazarus, S. (1990). Chapter 23. Emotion and Adaptation. In L.A. Pervin (Ed.). *Handbook of Personality: Theory and Research.* (pp. 609-637). New York: Guilford.

Southwick, S.M, Vythilingam, M & charney, D.S. (2005). The psychobiology of depression and resilience to stress: implications for prevention and treatment. *Annual Review of Clinical Psychology*, 1(255).

Van der Molen, H. T., Perreijn, S. & Van den Hout, M. A. (2010). *Klinische psychologie: Theorieën en psychopathologie.* Groningen: Noordhoff Uitgevers.

Weinberg, R. S. & Gould, D. (2011). *Foundations of sport and exercise psychology*. Champaign, IL: Human Kinetics.

Wintre, M.G & Vallance, D.D. (1994). A developmental sequence in the comprehension of emotions: Intensity, multiple emotions, and valence. *Developmental Psychology*, 30(4), 509-514.

Appendix

Theoretical segment figures

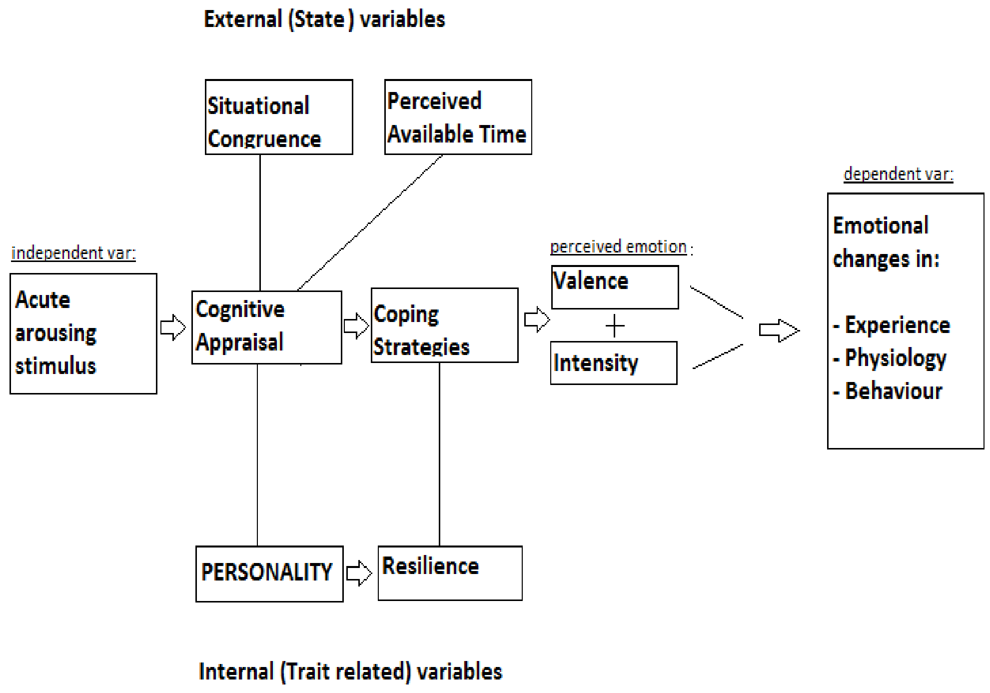

Figure 1: model of emotional activation

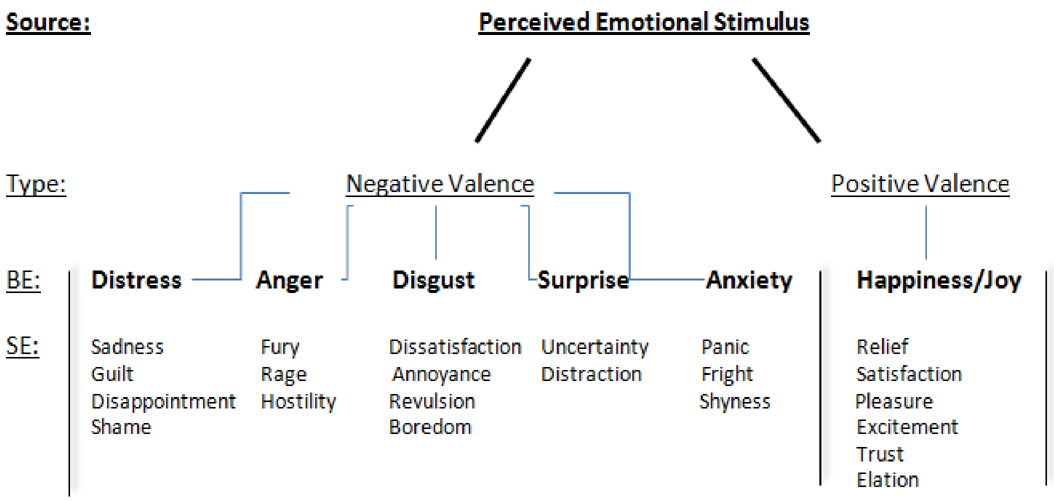

Figure 2: Valence classification

Data analysis output graphs

CHANGE SCORES FORMULAS:

SCR_change_0 = neutraal_4 – pre_baseline_4

SCR_change_1 = voorbereiding – neutral_4

SCR_change_2 = zingen – voorbereiding

SCR_change_3 = post_baseline_2 – zingen

SCR_change_4 = post_baseline_4 – post_baseline_2

HR_change_0 = HR_neutraal_4 – HR_pre_baseline_4

HR_change_1 = HR_voorbereiding – HR_neutraal_4

HR_change_2 = HR_zingen – HR_voorbereiding

HR_change_3 = HR_post_baseline_2 – HR_zingen

HR_change_4 = HR_post_baseline_4 – HR_post_baseline_2

Selfreport_change_1 = Stress_preparing – Stress_start

Selfreport_change_2 = Stress_singing – Stress_preparing

Selfreport_change_3 = Stress_after_singing – Stress_singing

Selfreport_change_4 = Stress_end – Stress_after_singing

Formulas regarding how to calculate change scores for all phases

Total group data (whole respondent pool)

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
SCR_change_0	,904	119	,000
SCR_change_1	,941	119	,000
SCR_change_2	,960	119	,001
SCR_change_3	,963	119	,002
SCR_change_4	,923	119	,000
HR_change_0	,879	119	,000
HR_change_1	,944	119	,000
HR_change_2	,962	119	,002
HR_change_3	,966	119	,004
HR_change_4	,820	119	,000
Selfreport_change_1	,961	119	,002
Selfreport_change_2	,951	119	,000
Selfreport_change_3	,891	119	,000
Selfreport_change_4	,955	119	,001

Normality analysis of whole participant pool

Correlations

	SCR_change_0	SCR_change_1	SCR_change_2	SCR_change_3	SCR_change_4
	,057	-,053	,065	-,090	,027
HR_change_0	**,524**	,556	,471	,319	,763
	126	126	126	126	126
	-,128	**,024**	-,014	,071	-,063
HR_change_1	,154	**,791**	,873	,428	,483
	126	**126**	126	126	126
	,103	-,095	**,027**	,009	,022
HR_change_2	,247	,285	**,764**	,919	,804
	128	128	**128**	128	128
	-,036	,010	-,060	**,011**	,061
HR_change_3	,689	,915	,502	**,901**	,495
	128	128	128	**128**	128
	-,078	,146	-,100	,110	**-,004**
HR_change_4	,383	,101	,263	,216	**,965**
	127	127	127	127	**127**

Automatic within system correlation graph of whole participant pool (using change scores)

Correlations

	HR_change_1	HR_change_2	HR_change_3	HR_change_4
	,227	-,024	-,055	-,066
Selfreport_change_1	**,012**	,789	,541	,469
	122	124	124	123
	-,160	**,128**	-,075	,013
Selfreport_change_2	,079	**,156**	,408	,888
	122	**124**	124	123
	-,069	-,171	**,243**	-,028
Selfreport_change_3	,449	,058	**,007**	,755
	122	124	**124**	123
	,007	-,104	,087	**,187**
Selfreport_change_4	,942	,252	,337	**,038**
	122	124	124	**123**

Change scores between system correlation graph including heartrate data (total group)

Correlations

	SCR_change_1	SCR_change_2	SCR_change_3	SCR_change_4
	,069	,123	-,112	,018
Selfreport_change_1	**,432**	,160	,200	,835
	133	133	133	133
	-,140	**-,076**	,163	-,084
Selfreport_change_2	,107	**,385**	,061	,334
	133	**133**	133	133
	-,108	,010	**,173**	-,053
Selfreport_change_3	,216	,910	**,047**	,546
	133	133	**133**	133
	,132	-,086	-,086	**,161**
Selfreport_change_4	,129	,327	,325	**,065**
	133	133	133	**133**

Change scores between system correlation graph including skin conductance data (total group)

GROUP 1 data: Experienced singing group (positive valence)

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
SCR_change_0	,928	61	,002
SCR_change_1	,956	61	,028
SCR_change_2	,959	61	,042
SCR_change_3	,952	61	,018
SCR_change_4	,927	61	,001
HR_change_0	,831	61	,000
HR_change_1	,940	61	,005
HR_change_2	,941	61	,005
HR_change_3	,935	61	,003
HR_change_4	,932	61	,002
Selfreport_change_1	,932	61	,002
Selfreport_change_2	,916	61	,000
Selfreport_change_3	,910	61	,000
Selfreport_change_4	,946	61	,009

Normality analysis of group 1

Correlations

	SCR_change_0	SCR_change_1	SCR_change_2	SCR_change_3	SCR_change_4
	,100	-,120	,011	-,009	,004
HR_change_0	**,417**	,329	,929	,939	,971
	68	68	68	68	68
	-,073	**,005**	,059	-,031	-,072
HR_change_1	,554	**,969**	,630	,804	,560
	68	**68**	68	68	68
	,081	-,022	**,005**	-,094	,007
HR_change_2	,506	,854	**,967**	,440	,957
	70	70	**70**	70	70
	-,118	,046	-,076	**,096**	,130
HR_change_3	,329	,703	,530	**,430**	,282
	70	70	70	**70**	70
	-,057	-,009	-,159	,274	**-,107**
HR_change_4	,644	,940	,191	,023	**,383**
	69	69	69	69	**69**

Automatic within system correlation graph of group 1 (using change scores)

Correlations

	HR_change_1	HR_change_2	HR_change_3	HR_change_4
	,278	-,030	-,008	-,087
Selfreport_change_1	**,026**	,810	,951	,490
	64	66	66	65
	-,135	**,098**	-,114	,129
Selfreport_change_2	,289	**,434**	,361	,305
	64	**66**	66	65
	-,123	-,356	**,393**	,086
Selfreport_change_3	,333	,003	**,001**	,497
	64	66	**66**	65
	-,127	,002	,108	**,118**
Selfreport_change_4	,317	,988	,389	**,349**
	64	66	66	**65**

Change scores between system correlation graph including heartrate data (group 1)

Correlations

	SCR_change_1	SCR_change_2	SCR_change_3	SCR_change_4
	,090	,157	-,154	,111
Selfreport_change_1	**,451**	,184	,193	,348
	73	73	73	73
	-,025	**-,223**	,232	-,211
Selfreport_change_2	,837	**,058**	,048	,073
	73	**73**	73	73
	-,144	-,069	**,252**	,077
Selfreport_change_3	,224	,564	**,031**	,519
	73	73	**73**	73
	-,013	,106	-,148	**,162**
Selfreport_change_4	,912	,371	,212	**,171**
	73	73	73	**73**

Change scores between system correlation graph including skin conductance data (group 1)

GROUP 2 data: Unexperienced singing group (negative valence)

Tests of Normality

	Shapiro-Wilk		
	Statistic	df	Sig.
SCR_change_0	,865	56	,000
SCR_change_1	,884	56	,000
SCR_change_2	,950	56	,020
SCR_change_3	,952	56	,026
SCR_change_4	,899	56	,000
HR_change_0	,902	56	,000
HR_change_1	,920	56	,001
HR_change_2	,969	56	,153
HR_change_3	,983	56	,617
HR_change_4	,735	56	,000
Selfreport_change_1	,962	56	,078
Selfreport_change_2	,949	56	,019
Selfreport_change_3	,847	56	,000
Selfreport_change_4	,947	56	,016

Normality analysis of group 2

Correlations

	SCR_change_0	SCR_change_1	SCR_change_2	SCR_change_3	SCR_change_4
	-,074	,078	,150	-,213	-,003
HR_change_0	**,587**	,568	,269	,116	,984
	56	56	56	56	56
	-,145	**-,022**	-,046	,143	,011
HR_change_1	,286	**,873**	,738	,295	,938
	56	**56**	56	56	56
	,155	-,192	**,096**	,187	,067
HR_change_2	,255	,156	**,482**	,167	,624
	56	56	**56**	56	56
	,055	,005	-,108	**-,054**	-,098
HR_change_3	,689	,972	,426	**,691**	,473
	56	56	56	**56**	56
	-,081	,275	-,055	-,123	**,137**
HR_change_4	,555	,040	,689	,367	**,316**
	56	56	56	56	**56**

Automatic within system correlation graph of group 2 (change scores)

Correlations

	HR_change_1	HR_change_2	HR_change_3	HR_change_4
	,125	-,034	-,088	-,083
Selfreport_change_1	**,358**	,803	,517	,544
	56	56	56	56
	-,168	**,184**	-,057	-,078
Selfreport_change_2	,216	**,175**	,675	,568
	56	**56**	56	56
	,006	,130	**,041**	-,125
Selfreport_change_3	,965	,341	**,765**	,357
	56	56	**56**	56
	,137	-,294	,108	**,271**
Selfreport_change_4	,315	,028	,430	**,043**
	56	56	56	**56**

Change scores between system correlation graph including heartrate data (group 2)

Correlations

	SCR_change_1	SCR_change_2	SCR_change_3	SCR_change_4
	,044	,023	-,060	-,034
Selfreport_change_1	**,743**	,863	,655	,800
	58	58	58	58
	-,247	**,185**	,072	,008
Selfreport_change_2	,062	**,164**	,593	,953
	58	**58**	58	58
	-,213	,055	**,167**	-,124
Selfreport_change_3	,108	,682	**,209**	,355
	58	58	**58**	58
	,368	-,323	-,087	**,106**
Selfreport_change_4	,004	,014	,518	**,430**
	58	58	58	**58**

Change scores between system correlation graph including skin conductance data (group 2)

Results table(s)

Table 1. Outcome of between system coherence correlation analysis of heartrate and self-reported differences for the positive valence group over the phases of the research. Correlation values (r=) are displayed together with the corresponding significance (p=). Relevant values are those between corresponding timeframes (1-4).

Change score variables	**Heartrate changes 1**	**Heartrate changes 2**	**Heartrate changes 3**	**Heartrate changes 4**
Selfreport changes 1	*r*= ,278 *p*= ,026	*r*=-,030 *p*= ,810	r= -,008 p= ,951	r= -,087 p= ,490
Selfreport changes 2	*r*= -,135 *p*= ,333	*r*= ,098 *p*= ,434	r= -,114 p= ,361	r= ,129 p= ,305
Selfreport changes 3	*r*= -,123 *p*= ,026	*r*=-,356 *p*= ,003	r= ,393 p= ,001	r= -,086 p= , 497
Selfreport changes 4	*r*= -,127 *p*= ,317	*r*= 002 *p*=,988	r= ,108 p= ,389	r= -,118 p= , 349