Dissertation

Response Synchrony and Response Patterning of Psychophysiological Parameters in Emotion

Technische Universität Darmstadt

von

Marlene Dejá

2018

Response Synchrony and Response Patterning of Psychophysiological Parameters in Emotion

vom Fachbereich Humanwissenschaften

der Technischen Universität Darmstadt

zur Erlangung des Grades

Doctor rerum naturalium

(Dr. rer. nat.)

Dissertation

von Marlene Dejá

Erstgutachter: Prof. Wolfgang Ellermeier, Ph.D.

Zweitgutachter: Prof. Dr. phil. nat. Augustin Kelava

Drittgutachter: Dr.-Ing. Michael Muma

Viertgutachter: Prof. Dr. Bernhard Schmitz

Darmstadt 2018

Dejá, Marlene: Response Synchrony and Response Patterning of Psychophysiological Parameters in Emotion
Darmstadt, Technische Universität Darmstadt
Jahr der Veröffentlichung der Dissertation auf TUprints: 2018
Tag der mündlichen Prüfung: 14.06.2018

to my family

Acknowledgment

I am very thankful to my supervisor Prof. Augustin Kelava for his scientific and moral support, his guidance, his patience, his feedback, his encouragement, and that he did not give up on me. I am grateful to Prof. Wolfgang Ellermeier for his comments, his support, and the willingness to supervise this dissertation. Further, I would like to thank my third supervisor Dr. Michael Muma, who helped me through various stages of this dissertation. His notes, his feedback, his explanations, his grammar corrections, and his loyalty helped me throughout this long and sometimes frustrating time.

Additionally, I would like to thank Johannes Weise and Stefan Richter, who helped me with the analysis of the data, who explained to me complex evaluation methods and solved technical difficulties throughout the data sampling process. Further, I would like to thank my friends David Hoppe and Holger Brandt, who made me a better R-user and who listened to my complaints more than once. I thank my friend Lisa for her grammar corrections and my friends Steffi, Jule, Annette, Isabel, Daniel, Sevda, Yvonne, and Kati for always being there for me.

I am also very grateful to my husband, who was happy for me and suffered with me throughout this process, who always supported me, and never stopped believing in me. I am also very thankful to my parents, my brother, and Svenja who supported and encouraged me during this time. Further, I would like to thank my parents-in-law (especially Claus for his grammar corrections) who morally supported me and regularly took care of my son, while I was sitting in front of my computer.

Wirkliches Neuland in einer Wissenschaft kann wohl nur gewonnen werden, wenn man an einer entscheidenden Stelle bereit ist, den Grund zu verlassen, auf dem die bisherige Wissenschaft ruht, und gewissermaßen ins Leere zu springen.

Werner Heisenberg (1901 - 1976)

Abstract

An emotional experience is associated with changes in behavior (e.g., facial expression), physiological parameters (e.g., increased heart rate), and subjective experience (e.g., feeling anxious). The different response parameters are said to be changing synchronously during an emotion in order to ensure an optimal reaction to the triggering stimulus (e.g., to flee from a bear; Ekman, 1992; Levenson, 1994). The common simultaneous change over time is referred to as response synchrony (Bulteel et al., 2014). According to the basic emotion approaches, synchrony is regarded as an essential component of emotional experience (Ekman, 1992). In empirical studies, however, the results concerning the synchrony of different response parameters are diverse (Hollenstein & Lanteigne, 2014). The lack of empirical support may be due to the complex multivariate and non-stationary data structure which have a large effect on methods that make over-simplifying assumptions. For instance, previous approaches for quantifying synchrony disregarded the non-stationarity of the data, that is the fluctuation of mean and variance over time, and analyzed data only on an interindividual level (e.g., averaging over several individuals). The possibility to describe synchrony in the course of time or to provide evidence of synchrony in single individuals is thus not given. On the other hand, there are theoretical approaches such as psychological construction approaches that question the necessity of synchrony for an emotional experience (e.g., Barrett, 2006a; Cunningham, Dunfield, & Stillman, 2013). Therefore, one aim of this doctoral thesis was to develop and to apply a new approach to quantify physiological synchrony.

Related to the question of the synchronous change of physiological parameters is the question whether individuals or even different emotional states can be correctly classified based on changes in physiological parameters. Here, the focus lies on the change of physiological parameters to a specific pattern depending on the evoked emotion which is described by the term response patterning (Bulteel et al., 2014). The majority of the studies collect data to classify individuals or emotions

only on one measurement occasion. Such a procedure neglects the daily variations and intraindividual changes of physiological data at different times and tends to overestimate the accuracy of the classification (R. A. Calvo, Brown, & Scheding, 2009; Picard, Vyzas, & Healey, 2001). For this reason, another aim of this doctoral thesis was the classification of individuals and emotions when data were collected at two different measurement occasions.

This cumulus contains three Manuscripts. The aim of Manuscript A ($N = 58$) was to develop a new time-frequency-based approach to quantify synchrony of physiological parameters on a latent level. Using the new approach, multivariate and non-stationary time series can be analyzed on an intraindividual level. The quantification of synchrony consists of two steps. In a first step, time-varying bivariate coherences of two physiological signals (e.g., electrocardiogram and electrodermal activity) are calculated. Due to the joint time-frequency-based approach, the non-stationarity of the data is taken into account. In a second step, these bivariate coherences are used in a state-space model as manifest indicator variables to form a latent synchrony variable at time t. The synchrony measure can take values between 0 and 1, where 0 means that the manifest coherences variables are completely uncorrelated and 1 means that they change synchronously. The results showed that the overall physiological synchrony variable was close to 1 in some parts of the film clip which was rated as more funny. Further, a high interindividual variability in the synchrony of physiological parameters was found. Compared to the network approach of Hsieh et al. (2011), the new method is capable of mapping the time course of physiological synchrony and revealing inter- and intraindividual differences. The network approach only returned results that counted for the entire sample under the assumption of stationarity and did not allow for individual variability.

The aim of Manuscript B ($N = 42$) was the further application of the newly developed approach for the quantification of physiological synchrony. The research question if the synchrony of physiological parameters during the emotional experience of disgust is higher than during a neutral emotional state was addressed. Further, the interindividual variability and the correlation between the subjective intensity level of disgust and the physiological synchrony were investigated. For this, participants were shown neutral and disgusting pictures. The results showed that synchrony was significantly higher shortly after showing a disgusting picture as compared to shortly after showing a neutral picture. At the same time, there were large interindividual differences in the temporal course of synchrony. Further, physiological synchrony started to increase before the actual picture was shown which raises the question, to what extent an orienting response can be responsible for the changes in physiological synchrony. The subjective intensity rating of disgust was measured continuously with a rating dial. It was at a maximum when physiological synchrony had already decreased back to the baseline level. A possible explanation could be motor and cognitive processes which are necessary for turning the rating dial.

The aim of Manuscript C ($N = 36$) was the classification of individuals and emotions by means of peripheral physiological data. In contrast to many previous studies, data were collected on two measurement occasions with a time interval of six weeks between them. Two well-established methods were applied as classifiers (k-nearest neighbors (KNN) and support vector machines (SVM)) that both take into account the nonlinear separability of the features that were extracted from the data. Pictures and film clips were used to induce fear. Fear could be better differentiated from a neutral state when film clips (77.50% KNN; 81.90% SVM) instead of pictures (64.40% KNN; 66.20% SVM) were used as induction method. Further, initial attempts were made to classify different levels of fear and to compare them with continuous ratings of fear. On a descriptive level, a connection between the classification and the subjective rating could be shown. In addition to the emotion classification task, individuals were classified using features from the electrocardiogram signal. In terms of classifying individuals, the correct classification rate showed a clear decline from 54.53% to 23.16% using the KNN and from 56.70% to 26.93% using the SVM when the training and testing data were from two different measurement occasions. This result demonstrates the high intraindividual variability of physiological data. However, compared to previous results, the classification rates were rather low which could be related to the emotion induction on a rather low intensity level.

In summary, in this doctoral thesis the synchrony of physiological parameters during an emotion is examined. Further, this thesis investigates to what extent physiological changes can be used to distinguish a given emotion from a neutral state. In a general discussion, the results are compared with the prevailing emotion theories. In conclusion, the results of this thesis show that physiological synchrony during an emotion exists with great interindividual differences. The results suggest that future studies on physiological parameters should not be evaluated on an interindividual, aggregated level, but rather consider intraindividual processes.

Zusammenfassung

Eine Emotion soll mit Veränderungen im Verhalten (z.B. Gesichtsausdruck), in physiologischen Parametern (z.B. erhöhte Herzfrequenz) und im subjektiven Erleben (z.B. Angst zu empfinden) einhergehen. Die unterschiedlichen Reaktionsparameter sollen sich während einer Emotion synchron verändern, um eine optimale Reaktion gegenüber dem auslösenden Stimulus (z.B. von einem Bären fliehen) zu gewährleisten (e.g., Ekman, 1992; Levenson, 1994). Die gemeinsame, simultane Veränderung über die Zeit wird auch Response Synchronität genannt (Bulteel et al., 2014). Synchronität wird nach den Basic Emotion Theorien als essentieller Bestandteil des emotionalen Erlebens angesehen (Ekman, 1992). In der Empirie sind die Ergebnisse bezüglich Synchronität verschiedener Reaktionsparameter jedoch uneinheitlich (Hollenstein & Lanteigne, 2014). Die mangelnde empirische Bestätigung von Synchronität kann an den komplexen multivariaten und nicht-stationären Daten liegen, die einen großen Effekt auf Methoden haben, die die Annahmen über die vorhandene Datenstruktur vereinfachen. Zum Beispiel, bisherige Methoden zur Quantifizierung von Synchronität vernachlässigen die nicht vorhandene Stationarität der Daten, sprich die Schwankungen von Mittelwert und Varianz über die Zeit und analysieren die Daten nur auf einer interindividuellen (d.h. gemittelt über mehrere Personen) Ebene. Die Möglichkeit, Synchronität im Zeitverlauf zu beschreiben oder für einzelne Personen nachzuweisen, ist so nicht gegeben. Auf der anderen Seite gibt es theoretische Ansätze, wie die psychologisch-konstruktivistischen Emotionstheorien, die die Notwendigkeit der Synchronität für ein emotionales Erleben in Frage stellen (e.g., Barrett, 2006a; Cunningham et al., 2013). Ein Ziel dieser Doktorarbeit war daher die Entwicklung und Anwendung einer neuen Methode zur Quantifizierung der Synchronität von physiologischen Parametern.

Verwandt mit der Frage der synchronen Veränderung physiologischer Parameter ist die Frage, ob Personen oder auch verschiedene emotionale Zustände aufgrund der Veränderung von physiologischen Parametern korrekt klassifiziert werden kön-

nen. Dabei liegt der Fokus auf den Veränderungen der physiologischen Parameter zu einem bestimmten Muster, in Abhängigkeit von der hervorgerufenen Emotion. Dies wird auch als Response Muster bezeichnet (Bulteel et al., 2014). Die Mehrzahl der Studien erhebt Daten zur Klassifizierung von Personen oder Emotionen nur an einem Messzeitpunkt. Diese Vorgehensweise vernachlässigt Tagesschwankungen sowie intraindividuelle Veränderungen physiologischer Daten zu unterschiedlichen Zeitpunkten und führt in der Regel zu einer Überschätzung der Klassifikationsrate (R. A. Calvo et al., 2009; Picard et al., 2001). Aus diesem Grund war ein weiteres Ziel dieser Doktorarbeit die Klassifizierung von Personen und Emotionen, wenn Daten zu zwei verschiedenen Messzeitpunkten erhoben wurden.

In diesem Kumulus sind drei Manuskripte enthalten. Ziel von Manuskript A ($N = 58$) war die Entwicklung einer neuen Zeit-Frequenz basierten Methode auf latenter Ebene zur Quantifizierung von Synchronität physiologischer Parameter. Mit Hilfe der neuen Methode können multivariate und nichtstationäre Zeitreihen auf einem intraindividuellen Level analysiert werden. Die Quantifizierung der Synchronität besteht aus zwei Schritten. Im ersten Schritt werden zeitlich variierende, bivariate Kohärenzen von jeweils zwei physiologischen Signalen (z.B. Elektrokardiogramm und Hautleitfähigkeit) berechnet. Aufgrund des Zeit-Frequenz basierten Ansatzes werden dabei die Nichtstationarität der Daten berücksichtigt. In einem zweiten Schritt werden diese bivariaten Kohärenzen in einem State-Space Modell als manifeste Indikatorvariablen zur Bildung einer latenten Synchronität zum Zeitpunkt t verwendet. Das Synchronitätsmaß kann Werte zwischen 0 und 1 annehmen, wobei 0 bedeutet, dass die manifesten Kohärenz Variablen unkorreliert sind und 1 bedeutet, dass sie sich synchron verändern. Die Ergebnisse zeigten, dass die physiologische Synchronität für den Film, der als "lustiger" bewertet wurde, nahezu 1 war. Ferner, wurde eine hohe interindividuelle Variabilität in der Synchronität physiologischer Parameter gefunden. Im Vergleich zu dem Netzwerkansatz von Hsieh et al. (2011), ist die neue Methode in der Lage den Zeitverlauf von physiologischer Synchronität sowie inter- und intraindividuelle Unterschiede aufzuzeigen. Die Ergebnisse des Netzwerkansatzes galten nur für die gesamte Stichprobe unter der Annahme von Stationarität und bildeten keine individuelle Variabilität ab.

Ziel von Manuskript B ($N = 42$) war die weitere Anwendung der neu entwickelten Methode zur Quantifizierung physiologischer Synchronität. Dabei sollte die Hypothese geprüft werden, ob die Synchronität physiologischer Parameter während des emotionalen Erlebens von Ekel höher ist als während eines neutralen emotionalen Zustandes. Des Weiteren wurden die interindividuelle Variabilität und der Zusammenhang zwischen der subjektiven Ekelintensität und der physiologischen Synchronität untersucht. Hierfür schauten die Probanden neutrale und eklige Bilder an. Die Ergebnisse zeigten, dass die Synchronität kurz nachdem ein ekliges Bild gezeigt wurde signifikant höher war als kurz nach einem neutralen Bild. Gleichzeitig gab es große interindividuelle Unterschiede im zeitlichen Verlauf. Ferner, fing die physiologische Synchronität an zu steigen, bevor das eigentliche Bild gezeigt wurde. Die subjektive Ekelintensität wurde kontinuierlich mit einem Regler gemessen und war maximal als die physiologische Synchronität bereits wieder auf das Grundniveau gesunken war. Eine mögliche Erklärung könnten motorische und kognitive Prozesse sein, die notwendig sind um den Regler zu drehen.

Ziel von Manuskript C ($N = 36$) war die Klassifizierung von Personen und Emotionen mit Hilfe von physiologischen Daten. Im Vergleich zu vielen bisherigen Studien wurden anwendungsnah Daten zu zwei verschiedenen Messzeitpunkten mit einem Zeitabstand von sechs Wochen erhoben. Als Klassifizierer wurden zwei etablierte Methoden (k-nearest neighbours (KNN) und support vector machines (SVM)) angewendet, die die nichtlineare Trennbarkeit der Features, die aus den Daten extrahiert wurden, berücksichtigen. Bilder und Filmausschnitte wurden verwendet um Angst zu induzieren. Angst konnte besser von einem neutralen Zustand differenziert werden, wenn Filme (77.50% KNN; 81.90% SVM) anstelle von Bildern (64.40% KNN; 66.20% SVM) als Induktionsmethode verwendet wurden. Des Weiteren wurden erste Versuche unternommen, unterschiedliche Intensitäten von Angst zu klassifizieren und diese mit kontinuierlichen subjektiven Ratings der Angstintensität zu vergleichen. Ein Zusammenhang konnte auf der deskriptiven Ebene gezeigt werden. Zusätzlich zu der Klassifizierung von Emotionen wurden Personen mit Hilfe der Features aus dem Elektrokardiogramm klassifiziert. In Bezug auf die Klassifizierung von Individuen sank die Klassifikationsrate von 54.53% zu 23.16% beim KNN und von 56.70% zu 26.93% beim SVM, wenn die Training- und Testdaten von zwei verschiedenen Messzeitpunkten stammten. Dieses Ergebnis verdeutlich die hohe intraindividuelle Variabilität physiologischer Daten. Dennoch sind die hier berichteten Klassifikationsraten im Vergleich zu bisherigen Studien niedrig. Dies könnte mit der Emotionsinduktion auf einem niedrigen Intensitätsniveau in Zusammenhang stehen.

Zusammenfassend wird in dieser Doktorarbeit die Synchronität von sich verändernden physiologischen Parametern während einer Emotion untersucht. Des Weiteren wird geschaut, inwieweit sich Veränderungen von physiologischen Daten dazu eignen, eine Emotionen von einem neutralen Zustand zu unterscheiden. In der allgemeinen Diskussion werden die Ergebnisse in Bezug zu den vorherrschenden Emotionstheorien gesetzt. Im Allgemeinen zeigen die Ergebnisse dieser Arbeit, dass physiologische Synchronität während einer Emotion mit großen interindividuellen Unterschieden existiert. Die Ergebnisse weisen darauf hin zukünftige Studien in Bezug auf physiologische Parameter nicht auf einem interindividuellen, aggregierten Level auszuwerten, sondern intraindividuelle Prozesse zu beachten.

Contents

List of Figures

List of Tables

1. Theoretical Background

1.1. Synchrony of Response Parameters

There is still no theoretical and empirical consensus on the question whether - during an emotion - different response parameters undergo correlated changes (Barrett, 2006a; Ekman, 1992). However, there is a broad consensus that emotions go along with changes in the behavioral (e.g., facial expression), the physiological (e.g., heart rate), and the subjective (e.g., self-reported emotional states) response modality (e.g., Mauss & Robinson, 2009). Most studies have used different terms to describe the simultaneous change of behavior, physiology, and subjective experience: coherence (e.g., Dan-Glauser & Gross, 2013; Hsieh et al., 2011; Mauss, Levenson, McCarter, Wilhelm, & Gross, 2005), concordance (e.g., Butler, Gross, & Barnard, 2014; Friedman, Stephens, & Thayer, 2014; Hollenstein & Lanteigne, 2014), synchrony (e.g., Grandjean, Sander, & Scherer, 2008; Scherer, 2009), and the organization of response systems (e.g., Levenson, 1994). Throughout this cumulus generally the term synchrony or response synchrony will be used to describe the common change between these different response modalities that are linearly related. Studies that only examine the simultaneous change within one response system such as facial expressions (e.g., Gentsch, Grandjean, & Scherer, 2014) or physiological changes (e.g., Bulteel et al., 2014) just evolved recently. This thesis concentrates on the synchrony of physiological changes. The term physiological synchrony will be used to refer to the synchronous changes only within the physiological response modality.

Further, it is important to distinguish between response *patterning* and response *synchrony* (see Bulteel et al., 2014). On the one hand, response patterning refers to the change in one or all modalities to a specific pattern depending on the evoked emotion (Bulteel et al., 2014). This refers to the direction of change in various parameters during an emotional episode (e.g., heart rate and skin conductance increase during fear, Kreibig, 2010). Discrete recurring response patterns for different emotions are very important for the ability to classify different emotional states in humans. On the other hand, response synchrony describes the increased covariation of different parameters over time (Bulteel et al., 2014). Hence, the direction of the change is not important, the focus is on correlated changes within a relatively short time window. The first part of the cumulus deals with response synchrony, how it can be quantified and what theoretical assumptions are behind it. The second part deals with response patterning in the form of the classification of emotions.

In the past 20 years, two theoretical approaches have determined the scientific study of emotions. In the following sections, both approaches are introduced and their core statements regarding the synchrony of response parameters are presented. First, the main assumptions of the basic emotion approaches are described. Subsequently, the main points of the psychological construction theories and their criticism of the basic emotion approaches are presented. Finally, both theoretical assumptions are placed in relation to the classification of emotions. A link between the synchrony and classification of emotions on the basis of physiological data is established.

1.1.1. Basic emotion approaches

The basic emotion approaches influenced the scientific models, methodologies, and experimental settings over the past 50 years. According to the basic emotion approaches (e.g., Ekman & Cordaro, 2011; Izard, 2011; Levenson, 2011; Panksepp & Watt, 2011; Tomkins, 1962), basic emotions fulfill an evolutionary function and they are distinctive emotional categories that meet certain criteria. The idea that the expression of distinct emotional states is a product of human evolution can be traced back to Darwin (1872/1965). Basic emotion approaches assume that emotions have direct causal influence on motivation and behavior in human beings (Tracy & Randles, 2011). This is especially the case, if the elicitors of emotions are related to an evolutionary basis (e.g, a bear appears during a walk), if they appear suddenly, and if they are intense (e.g., Izard, 2011; Levenson, 2011). The number of basic emotions varied, depending on the approach and the respective underlying criteria between six to eight basic emotions (see Tracy & Randles, 2011, for an overview). Recently, Ekman (2016) reported in his survey that the majority of the scientific community agreed upon five empirically established discrete emotions: anger, fear, disgust, sadness, and happiness.

The list of criteria for basic emotions is similar to the different approaches, so the criteria of Ekman and Cordaro (2011) and Levenson (2011) are discussed as representing typical examples. Ekman and Cordaro (2011) postulated amongst others the following criteria: distinctive universal signals, distinctive physiological pattern, automatic appraisal, distinctive universals in antecedent events, can be of brief duration, presence in other primates, capable of quick onset, unbidden occurrence, and distinctive subjective experience. Levenson (2011) condensed these criteria to three main criteria: distinctness in behavior, expression, and physiology; hard-wiredness (basic emotions can not be learned, but are predisposed from birth); functionality (basic emotions should solve problems which are critical to species survival and thriving).

Further, basic emotion approaches assume that every basic emotion can be assigned to a particular neural circuit in the brain (Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962). These circuits are supposed to be installed by evolution, since they ensure the survival of the species. The specific circuit is triggered through certain stimuli. Once activated it leads to specific action tendencies, specific responses, and specific emotional experiences (see Moors, 2009, for an overview). The synchronous response of different response parameters during an emotional experience (next to a distinctive facial expression, distinctive physiological patterns, and universality) is seen as indirect evidence for the existence of a neural circuit (Ekman, 1992). Hence, the synchrony of response parameters plays an important role regarding the empirical evidence of the basic emotion assumptions. Likewise distinctive physiological patterns are not only the indirect evidence for distinctive neural circuits in the brain but also lay the foundation for the possibility to classify different emotional states.

Some of the so-called appraisal approaches share similar ideas, with the difference that the emotions are not evoked through a stimulus itself, but rather through the person's appraisal of the stimulus (e.g., Lazarus, 1991; Roseman, Spindel, & Jose, 1990). Figure 1.1 depicts the assumptions about the emergence of an emotion according to the basic emotion approaches.

Functionality of synchrony in response parameters.

The development of an emotion is described by the basic emotion approaches as follows: A stimulus elicits an emotion, which, by activating certain brain circuits, causes emotion-specific changes in multiple response modalities (e.g., physiological

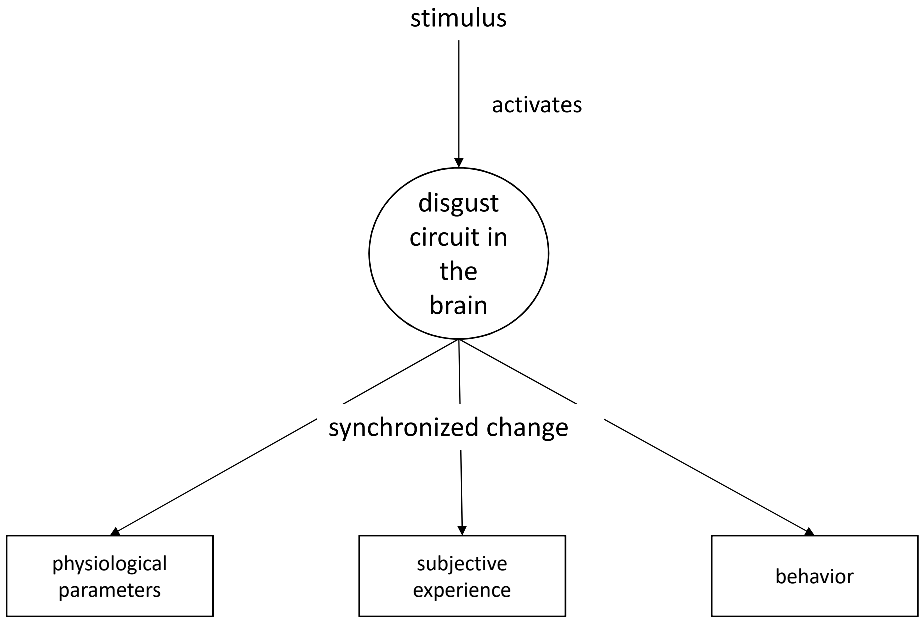


Figure 1.1. According to the basic emotion approaches (e.g., Ekman, 1992), a stimulus activates an emotion specific circuit in the brain, which leads to a synchronized change in physiological parameters, behavior, and subjective experience.

parameters, behavior, and subjective experience; Ekman & Cordaro, 2011; Izard, 2011; Panksepp & Watt, 2011). These changes ought to be synchronous and similar across individuals (Ekman, 1992). As synchronous responses between the different emotional response patterns serve as an optimal preparation towards environmental demands (e.g., fight or flight in dangerous situations), synchrony should increase with the intensity of an emotional stimulus (e.g., Davidson, 1992; Hodgson & Rachman, 1974; Levenson, 1994; Rosenberg & Ekman, 1994). Synchrony between or within the different response systems should be low if the emotion is weak, as there is no need for a physiologically/behavioral preparation (Hodgson & Rachman, 1974; Levenson, Ekman, & Friesen, 1990; Mauss et al., 2005).

For example, an emotion-eliciting situation, such as a contaminated toilet, would activate the disgust circuit, which would cause the synchronous activation of distinct response patterns of physiological (e.g., decreased heart rate), behavioral (e.g., wrinkling one's nose), and subjective parameters (e.g., feeling nauseated; see Figure 1.1). Thereby, the level of activation should correspond to the intensity level of disgust (e.g., Coan, 2010; Davidson, 1992; Hodgson & Rachman, 1974; Levenson, 1994; Rosenberg & Ekman, 1994). Summing up, from this point of view, the synchronous change between or within a response modality is seen as a core element of an emotional experience and serves as a biological function in terms of optimal preparation for action (Levenson, 1994; Rosenberg & Ekman, 1994).

1.1.2. Psychological construction approaches

Over the years the criticism of the basic emotion approaches has increased. In addition to theoretical considerations, the inconsistency and the lack of reproducibility of the scientific results were criticized and considered as insufficient evidence for the basic emotion approaches (Barrett, 2006a; Ortony & Turner, 1990). While Darwin, Tomkins, and Ekman saw an evolutionary function of emotions, interindividual variability was not always accounted for in their assumptions (Ekman, 1992). Interindividual variability indicates that individuals in the same situation may differ in the way they react. Intraindividual variability describes that the same individual in the same situation at a different time may react differently (see also Cacioppo et al., 1992). Therefore, James' (1884) ideas that emotional experiences are self perceptions (interpretation of information in the body) and that physiological change precedes the emotion instead of being a consequence of it, gained in importance. These ideas were picked up and developed further by representatives of psychological construction approaches (e.g., Barrett, 2006a; Cunningham et al., 2013; Ortony, Clore, & Collins, 1988; Russell, 2009). However, before the psychological construction approaches will be introduced in more detail, the profound criticism on the basic emotion approaches is presented.

1. Studies have reported distinctive cortical areas for basic emotions (e.g., Kassam, Markey, Cherkassky, Loewenstein, & Just, 2013; Kragel & LaBar, 2015; Saarimäki et al., 2016). However, others contradict these results and claim that empirical evidence for emotion-specific circuits in the brain is missing (e.g., Clark-Polner, Johnson, & Barrett, 2017; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Murphy, Nimmo-Smith, & Lawrence, 2003; Touroutoglou, Lindquist, Dickerson, & Barrett, 2015). Barrett (2017b) pointed out that the respective distinctive reaction pattern in the brain reported by some studies can not be reproduced in other studies.

2. Studies have reported specific physiological patterns for basic emotions (e.g.,

Kragel & LaBar, 2013; Kreibig, 2010; Stephens, Christie, & Friedman, 2010). Although the same elicitation method, population, and experimental methods are used, some of the results can not be replicated (Barrett, 2017b).

3. Studies have shown that there is a wide heterogeneity of emotional categories between cultures (e.g., Russell, 1991) and individuals (e.g., Ceulemans, Kuppens, & van Mechelen, 2012; Kuppens, van Mechelen, Smits, de Boeck, & Ceulemans, 2007; Nezlek, Vansteelandt, van Mechelen, & Kuppens, 2008), although homogeneity was originally expected by basic emotion approaches (Ekman, 1992). The borders between different emotions do not appear to be all that well-defined (Russell, 2009).

4. Response synchrony is assumed by the basic emotion approaches (as implicit evidence for an existing neural circuit; Ekman, 1992). However, empirical evidence regarding response synchrony is almost non-existent (Davidson, 1978; Friedman et al., 2014; Lazarus, Speisman, & Mordkoff, 1963; Mauss, Wilhelm, & Gross, 2004; Reisenzein, 2000) or exhibits great interindividual differences (Bulteel et al., 2014; Hsieh et al., 2011).

These major criticisms of the basic emotion approaches made the assumptions of the psychological construction approaches more important. In that framework, rather than postulating an affective program that has been installed in the course of evolution, an emotion is thought to emerge through the process of evaluating incoming contextual, bodily, and cognitive information. There is no fixed number of distinct emotional categories; instead, emotions have some sort of *core affect*, such as valence (e.g., Clore & Ortony, 2013; Cunningham et al., 2013), arousal (e.g., Schachter & Singer, 1962), or valence and arousal (e.g., Lindquist, Siegel, Quigley, & Barrett, 2013; Russell, 2003). According to Russell and Barrett (1999) the core affect of valence and arousal is also associated with distinctive neural circuits in the brain. However, compared to the evolutionary affect program of the basic emotion approaches, the situation, personal history, and experience all shape the affective core which is then labeled, for instance, disgust, thereby allowing for inter- and intraindividual variability (Barrett, 2013; Coan, 2010). Figure 1.2 depicts the assumptions about the emergence of an emotion according to the psychological construction approaches.

Functionality of synchrony in response parameters.

According to the psychological construction approaches, the synchrony of reaction parameters is not essential for the experience of an emotion (Barrett, 2006a; Clore & Ortony, 2013; Coan, 2010). One reason is the missing assumption of an emotion-specific circuit in the brain. Instead, psychological construction approaches place a special emphasis on physiological change (e.g., Clore & Ortony, 2013; Cunningham et al., 2013; Lindquist et al., 2013). For an emotion to emerge, a psychologically relevant situation causes or influences some sort of physiological change which is made psychologically meaningful through continuously reinterpreting one's sensation in the light of personal goals, intentions, and values (Barrett, 2013; Coan, 2010). Therefore, physiological change, such as decreasing heart rate, does not result from the elicitation of disgust; but rather, the situation (e.g., contaminated toilet) causes the physiological change along with other reactions (e.g., wrinkling one's nose or avoidance behavior), which, through ongoing appraisal processes, results in disgust (see Figure 1.2). Furthermore, individuals in the same situation - but also the same individual in the same situation at a different time - can experience different degrees of disgust (Clore & Ortony, 2013; Coan, 2010; Lindquist et al., 2013). On the other hand, if the disgust circuit in the basic emotion approaches is activated on a lower level, all response parameters must be activated on a lower level as well (Coan, 2010; Ekman, 1992; Tomkins, 1962).

Summing up, response synchrony within or between response modalities is seen as core element by the basic emotion approaches. Here, synchrony is the direct consequence of an evoked emotion and provides evidence for a causal mechanism in the brain. On the other hand, psychological construction approaches do not see response synchrony as essential for having an emotional experience. From their perspective, changes in response modalities lead through cognitive processes to an emotional episode. In this dissertation, the assumptions of the basic emotion approaches are taken as underlying theoretical model, which influenced the design of the experiments and the stimulus material. The critics of the basic emotion approaches take the lack of empirical evidence as proof that response synchrony does not exist. However, the specifics of how synchrony is quantified, the experimental design, and the level of analysis also have a considerable influence on the possible detection of response synchrony. This dissertation tries to meet some of the methodological issues regarding the quantification of response synchrony in emotions (see the following Section 1.1.3).

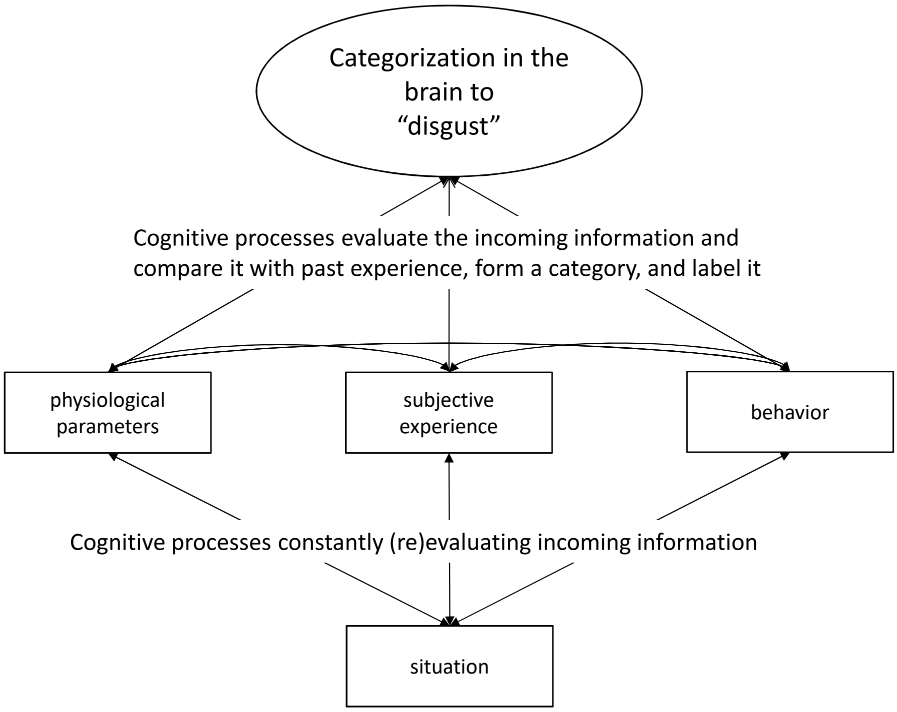


Figure 1.2. According to the psychological construction approaches (e.g., Barrett, 2006a), a situation leads to change in various response parameters. The changes do not have to happen at the same time. The changes are registered by the brain, evaluated by appraisal processes, compared with old experiences, and categorized. Throughout this process incoming information is considered.

1.1.3. Methodological Issues

In the literature, each study operationalized synchrony with different physiological, behavioral, and subjective parameters, quantification methods, and emotion induction methods. The influence of those methodological aspects on the analysis of response synchrony are discussed in the following section.

In general, the attempt of previous studies was to show the existence of an overall, aggregated increased synchrony during emotional episodes. These intentions have not been fulfilled, as many studies only found low or non existing response synchrony (e.g., Davidson, 1978; Friedman et al., 2014; Lazarus et al., 1963; Mauss et al., 2004). On the other hand, some studies support the idea that synchrony increased with the stimulus intensity (Mauss et al., 2005; Hsieh et al., 2011). There are several reasons for the low synchrony or the heterogenous findings. Theoretical considerations that assume response synchrony (Ekman, 1992) might not hold. However, if we assume that synchrony exists in principle, the apparent lack may be a result of methodological issues (see also Mauss et al., 2005). The first methodological difficulty to be mentioned here is the problem of inducing emotions of a certain type and intensity (e.g. Davidson, 1992; Rosenberg & Ekman, 1994). Low synchrony could be caused by inducing emotions on a low intensity level, for example, a film clip that is not funny enough. Further, the temporal development of response synchrony or of an emotional response in general is not well known (Bulteel et al., 2014). Hence, during a film sequence it is difficult to know in which segment of the data one should look for a synchronous response. Second, the multivariate, non-stationary data structure needs to be considered in order to detect response synchrony (Hollenstein & Lanteigne, 2014). For instance, only bivariate analysis or aggregation of data might affect the magnitude of the findings (e.g. Rosenberg & Ekman, 1994). A third methodological issue has been described by Cacioppo et al. (1992), where the authors emphasize that between- and within-perspectives on the data can lead to different conclusions regarding the relationship between variables.

Temporal dynamics of synchrony

The beginning, development over time, and end of an emotion are difficult to capture but crucial for the study of synchrony because simultaneous changes in various parameters in time series data are at the center of an analysis. Most studies have induced emotions with pictures (e.g., Bulteel et al., 2014; Dan-Glauser & Gross, 2013) or film clips (e.g., Hsieh et al., 2011; Mauss et al., 2005) in a laboratory setting. In most cases, the time period from the beginning to the end of an emotional experience has been seen as equal to the duration of an emotion-evoking stimulus (Hsieh et al., 2011; Friedman et al., 2014). Others chose different time lags such as 10 s (Mauss et al., 2005) or 8 s (Dan-Glauser & Gross, 2013). Nevertheless, the individual time frame of the emotional experience may deviate from the presented stimulus (Bulteel et al., 2014; Waugh & Schirillo, 2012). For instance, basic emotions are supposed to have a quick onset and duration (0 to 4 s; Ekman, 1984). Magnetic resonance imaging has detected changes in brain activity during disgust after 140 ms (Esslen, Pascual-Marqui, Hell, Kochi, & Lehmann, 2004; Hot & Sequeira, 2013). Likewise, facial expression changed within milliseconds after emotion induction (Ekman, 1992). However, the time course of physiological parameters or even physiological synchrony is largely unknown (Bulteel et al., 2014). Just recently, Bulteel et al. (2014) reported that physiological synchrony was highest when they presented a positive cue that signaled the valence of the upcoming picture. Thus, physiological change occurred solely due to the anticipation of an emotional cue. However, the possibility that synchrony functions as an orienting response to a stimulus and not as part of the emotional response itself could not be ruled out, either.

Quantification of response synchrony

In the past, response synchrony has been quantified using different approaches: cross-correlations (Dan-Glauser & Gross, 2013; Mauss et al., 2005; Sze, Gyurak, Yuan, & Levenson, 2010), ANOVA (Dan-Glauser & Gross, 2013), multilevel analysis (Butler et al., 2014), simultaneous regression analyses (Evers et al., 2014), DeCon (Bulteel et al., 2014), redundancy analysis (Friedman et al., 2014), hierarchical linear modeling (Hastings et al., 2009), or ROC-Analysis and Spearman Rank correlation (Hsieh et al., 2011). In the following, some of the approaches and results will be described in more detail.

Mauss et al. (2005) used time-lagged correlations to assess the covariation for each pair of measures (skin conductance and heart rate, heart rate and continuous subjective ratings, etc.) within an individual over time. They picked the maximum pairwise correlation within lags of -10 s to 10 s of each person. They showed that synchrony and intensity level of the emotional stimulus are positively associated for amusement, but not for sadness. Furthermore, they found high intraindividual variability, suggesting that people do not tend to react homogeneously. However, the applied method of Mauss et al. (2005) does not specify a single synchrony variable, therefore, only pairwise values and not one overall synchrony value can be reported.

Hsieh et al. (2011) developed a stochastic network configuration approach to quantify synchrony. They showed that synchrony of 15 physiological, behavioral and subjective variables was increased for more intense emotional stimuli (amusement and sadness compared to neutral states). They also discovered different response timings towards negative (sadness) and positive (amusement) stimuli, with the response coming and going quicker in the case of a positive emotion. The problem is that their results only account for the whole sample. However, they found that people differ in their synchrony response, questioning if results on the aggregated level are even appropriate.

Bulteel et al. (2014) analyzed physiological synchrony using a DeCon change point detection method. In contrast to the other approaches mentioned above, they analyzed data also on an intraindividual level. They found that during the presentation of a positive picture physiological parameters were on average more synchronized. However, while physiological synchrony was found on an interindividual level across participants, they were able to show large differences between individuals. For some individuals physiological synchrony rather decreased than increased while watching the positive slides. Further, they reported that the most consistent change in physiological synchrony was when the stimulus, which indicated the valence of the next picture, appeared, not the image itself. Their study was the first to make statements about the timing of physiological synchrony.

Level of analysis

Analyzing data with an interindividual approach provides information about synchrony on an aggregated level (e.g., individuals who report stronger emotional experience compared to other individuals would be expected to show a more synchronous physiological response; Mauss et al., 2005). On the other hand, analyzing data with an intraindividual approach provides information about the synchrony of an individual across time (e.g., synchrony would be expected to be higher during time periods where the individual reports more emotional experience compared to time periods where the individual reports less emotional experience; Mauss et al., 2005). Most studies have confirmed large differences between individuals in response synchrony (e.g., Bulteel et al., 2014; Hsieh et al., 2011; Mauss et al., 2005). These high variations often prevent that synchrony on an aggregated level can be found (Hollenstein & Lanteigne, 2014). In addition, different patterns of response styles for a given individual could lead to a lack of (global) synchrony across individuals, while in fact, it exists on the intraindividual level. Hence, the interindividual approach alone might not be the most adequate method for capturing how tightly response parameters are coupled during emotions (Cacioppo et al., 1992; Stemmler, 1992). Therefore, the analysis of synchrony within an individual across time is becoming a focus of current research (Bulteel et al., 2014). Instead of looking for an aggregated effect, the attention moved to intraindividual analysis, thus the analysis of the course of synchrony of one individual over a certain time period (Bulteel et al., 2014).

Summing up, the inconsistent results regarding the existence and level of response synchrony during an emotion might be due to theoretical considerations. However, methodological issues like emotion intensity, temporal development, underlying data structure, and level of analysis might have prevented that response synchrony has been reliably found in the past. The presented quantification approaches all have pitfalls: either they do not account for non-stationary data, they do not consider multivariate data, or they do not provide information on an intraindividual level, hence, course of synchrony for one individual across time. To adjust for these drawbacks, a new time-frequency quantification method was developed and applied which will be presented in Manuscript A and B.

1.2. Classification of Emotions

The theoretical approaches that have just been discussed above also play a role in the classification of different emotions. Instead of response synchrony, response patterning (the specific change of parameters for different emotions) is now the focus (see Section 1.1). Emotion or affect classification is especially important for human-computer interactions. One of the aims of affect detection is to respond to the user's affective state in an intelligent fashion to improve the usability of the respective device (e.g., computer games, smart phones, online-learning tutorials, cars, etc.; Cai & Lin, 2011; D'Mello, Craig, Witherspoon, McDaniel, & Graesser, 2008). However, affect detection may also play an important part in the interaction with patients that have difficulties to recognize emotions by themselves or in others, or are not able to communicate their own emotional state (e.g., patients with autism, elderly patients, or somatic patients; Liu, Conn, Sarkar, & Stone, 2008). D'Mello and Kory (2015), among others, point out that emotion classification is a difficult undertaking, since emotions are psychological constructs and not directly observable. Further, emotions are sensitive to the situational, social, individual, and cultural context (Barrett, 2006a; Russell, 2003). In addition, emotion classification is based on multivariate noise-sensitive data (e.g., physiological signals) or can be willingly influenced (e.g., facial expression or voice; Gu, Wong, & Tan, 2012; Picard et al., 2001).

Both of the theoretical approaches presented in this doctoral thesis are applied in the study of emotion classification. On the one hand, emotions are classified according to the assumptions of the basic emotion approaches, implying that discrete emotions cause unique and distinctive response patterns of physiological, behavioral, and subjective parameters (e.g., Kragel & LaBar, 2013; Stephens et al., 2010; Wei & Jia, 2016). On the other hand, some of the psychological construction approaches classify emotions according to the core affect of emotions, namely valence and arousal (Swangnetr & Kaber, 2013; Walter et al., 2013; Wiem & Lachiri, 2017). Some meta-analyses claim that emotion classification according to the assumptions of the basic emotion approaches provide better classification rates than the classification of valence and arousal (e.g., Kragel & LaBar, 2013; Lench, Flores, & Bench, 2011; Stephens et al., 2010). However, Barrett (2017a) criticized these studies. She claims that the unique emotion-specific patterns, which distinguished discrete emotions from each other in one study, are not replicated in other studies. Nevertheless, one aim of this thesis was to observe the temporal development of the subjective experience of fear and how different levels of fear can be accurately classified. Hence, the basic emotion approach was taken as the underlying theoretical model.

Apart from the theoretical approach that is used, D'Mello and Kory (2015) reported in their meta-analysis that 97% of all reviewed studies collected training and validation data on the same measurement occasion. This is a problem, since high inter- and intraindividual differences with regard to physiological data in emotions are reported (e.g., Kukolja, Popovic, Horvat, Kovac, & Cosic, 2014; Li, Xu, & Feng, 2016). Using only one measurement occasion leads to an overestimation of the correct classification rate and does not meet the demands of real world applications (Abdat, Maaoui, & Pruski, 2011; R. A. Calvo et al., 2009; Picard et al., 2001). One of the aims of this thesis was to classify an emotional state using two measurement occasions and therefore, be able to consider inter- and intraindividual variability of physiological data.

1.2.1. Classification of emotions and response synchrony

According to the basic emotion approaches, there should be synchrony between the individual response parameters, since they are triggered by a specific circuit in the brain (Ekman, 1992). Hence, the activation of the fear circuit should lead to changes in behavior (e.g., facial expression, voice, body posture), physiological parameters (e.g., heart rate, skin conductance, respiration), and subjective experience. These assumptions of the basic emotion approaches are very important for affect detection. If one wants to detect anger in a car driver, anger should be accompanied by specific, reliable changes in physiology and behavior. Hence, the classification rate should increase, if more than one modality (e.g., physiological data and behavior) is used to classify the emotional state. This assumption contains the connection between response synchrony and response patterning. Studies have shown that classification rates increased when additional modalities were included in the data used for classification (Abdat et al., 2011; Bailenson et al., 2008). These results support the basic emotion approaches. However, D'Mello and Kory (2015) argued in their meta-analysis that the additional multimodal effect was only modest, too modest to support the basic emotion assumptions and rather supporting the psychological construction approaches. With the presented quantification method of response synchrony in this thesis, it should be easier to answer these questions in the future.

1.3. Research Question

This doctoral thesis consists of three manuscripts - one of them published, one of them to be revised and resubmitted, and one of them submitted - each contains one experimental study. In the following section, the broad research questions, which derived from the theoretical considerations outlined above, are presented.

1. **Synchrony of physiological parameters in emotion**

   One of the major aims of this doctoral thesis is the presentation and application of a newly developed approach for the quantification of response synchrony. Thereby the focus lies on the quantification of physiological synchrony. The new approach quantifies one latent overall synchrony variable and takes the non-stationarity of the multivariate physiological data into account. One goal is to present results not only on an interindividual, aggregated level but also on an intraindividual level. Linked to this goal is the clarification of two theoretical aspects: First, can physiological synchrony for emotions like amusement or disgust be found with an appropriate quantification approach? Second, if it exists, is it similar for different individuals, how is the level and time course, and how is the time course related to the subjective experience? These questions are based on the two opposing theoretical approaches (basic emotion approaches and psychological construction approaches) which have different views on physiological response synchrony and its function. The research questions have a descriptive character since previous solutions for these problems do not exist. Therefore, there is little to no prior knowledge of the level, temporal course, and interindividual differences of physiological response synchrony. These research questions will be addressed in Manuscript A and B.

2. **Classification of fear using physiological parameters from multiple measurement occasions**
The first part of the dissertation deals with physiological response synchrony while the second part refers to response patterning and the classification of emotion based on physiological parameters. The aim is to see whether fear can be distinguished from a neutral emotional state when training and testing data do not derive from the same measurement occasion. Further, individuals are classified based on physiological data of two measurement occasions. Here, the influence of the number of measurement occasions as well as that of a larger sample size (more individuals) on the classification accuracy is investigated. Two prominent classifiers, k-nearest neighbors (KNN; Cover & Hart, 1967) and support vector machines (SVM; pioneered by Vapnik, 1998) are compared to each other as well as pictures with film clips in terms of the induction methods. The study design offers a larger sample size compared to previous research as well as two measurement occasions and hence, considers realistic conditions that better meet the requirements of real world applications. Third, on the basis of the physiological data, predictive probabilities for belonging to the fear class were estimated over time and compared with a continuous self-reported rating of fear. The goal is to compare the probability estimates with continuous self-reported data and to analyze, if a higher probability estimation for the fear class is accompanied by a higher subjective fear rating and vice versa. Prior studies only used retrospective subjective experience data for emotion intensity. These research questions will be addressed in Manuscript C.

2. Overview of the Manuscripts

2.1. Manuscript A: A new Approach for the Quantification of Synchrony of Multivariate Non-Stationary Psychophysiological Variables During Emotion Eliciting Stimuli[1]

2.1.1. Purpose of the study

The synchrony of response parameters is seen as a fundamental part of an emotional episode by various emotion theories (e.g., Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962). Others question whether the simultaneous change between, for example, physiological parameters and facial expression, is even necessary for an emotional episode (e.g., Barrett, 2013; Coan, 2010; Lewis, 2005). In the past, empirical findings regarding response synchrony during an emotional experience have been inconsistent. Only a few studies were able to find synchrony within or between response modalities (Bulteel et al., 2014; Hsieh et al., 2011). The lack of empirical evidence can have several reasons: First, the existence of response synchrony proposed by basic emotion theories might be false. Second, methodologies pertinent to an 'interindividual subjects design', like computing correlations across individuals, have been applied (e.g., Mauss et al., 2005). Aggregating data across individuals may disguise synchronous responses of sole individuals (Bulteel et al., 2014). Hence, beside the theoretical assumptions, synchrony might not have been found due to the methodological approach applied (Hollenstein & Lanteigne, 2014). Third, if analyzing synchrony of response parameters one is confronted with multivariate and non-stationary indicators where distributions of variables change

[1]Kelava, A., Muma, M., Deja, M., Dagdagan, J. Y., & Zoubir, A. M. (2015). A new approach for the quantification of synchrony of multivariate non-stationary psychophysiological variables during emotion eliciting stimuli. *Frontiers in Psychology, 5.* DOI: 10.3389/fpsyg.2014.01507.

over time (Butler et al., 2014). When the non-stationarity of the underlying data is not considered, information about the temporal development of synchrony is lost. Hence, new tools should be developed to adequately quantify response synchrony (Gross, 2010). The purpose of Manuscript A was to develop and apply a new time-frequency based latent variable approach to measure intra- as well as interindividual physiological synchrony of multivariate non-stationary time series.

2.1.2. Method

First, a time-frequency based approach was used to quantify bivariate coherences (Matz & Hlawatsch, 2000; Muma, Iskander, & Collins, 2010; White & Boashash, 1990), for example, the coherence of the ECG and respiratory signals. In this case, coherence describes the linear relation at a certain frequency between to signals. Time-frequency bivariate coherence account for the non-stationarity of the underlying physiological signals, hence the temporal variation of its mean and covariance function (Muma et al., 2010). In total, six pairwise coherence time series were calculated. Each coherence time series as well as the overall latent variable later on is bounded by 0 and 1. If the two signals are completely uncorrelated, their coherence equals zero. Second, a latent state-space model (Durbin & Koopman, 2001) was used to quantify one latent synchronized variable. The latent state-space model consists of a measurement model and a structural model. The basic idea of the state-space procedure is similar to a structural equation model. First, the measurement model uses the six pair-wise coherences as indicator variables and operationalizes a latent variable that represents an underlying state of the overall system-wise physiological synchrony at a certain point in time. It gives an overall latent physiological synchrony variable (latent state), for example, 100 ms after the start of the experiment. The same procedure is conducted 200 ms after the start of the experiment. In the structural model, the so resulting latent state vector is (auto-)regressed on a previous state vector. In this way, continuous predictions of latent states can be calculated, for example, by applying the Kalman filter (e.g. Grewal & Andrews, 2001; Shumway & Stoffer, 2006). The time series of the resulting latent states represents the course of the overall physiological synchrony. The approach was applied to a data set of 58 German male students who watched two funny film clips. Physiological measures were continuously sampled with the Biopac MP 150 System and the AcqKnowledge 4.2 Software (Biopac Systems Inc, 2011). The analysis was restricted to the electrocardiogram (ECG), the electrodermal activity

(EDA), and the respiration signal. Retrospective subjective data were collected after watching the film clip.

2.1.3. Results and discussion

In total, six pairwise coherence indicators were obtained using the ECG, EDA, and respiration signal. These six pair-wise coherences were used to build an overall latent synchrony variable. The results showed that individuals responded differently towards the funny film clips and had different temporal developments. For the film clip, which was rated as "more funny", physiological synchrony was close to 1 for some participants. This result shows that physiological synchrony can be found in empirical data and hence, that physiological synchrony does exist which goes along with the assumptions of the basic emotion approach. Manuscript A compared the new approach with the network approach of Hsieh et al. (2011). When the same data was analyzed with the network approach, a probability of 5.1% to achieve system-wise synchrony was calculated under the assumption of stationarity. This result counts for the whole sample and does not differentiate between individuals. Therefore, it neglects that the probability might differ for different participants at different points in time. On the other hand, the time-frequency approach is able to detect physiological synchrony for sole individuals over a certain time period. Further, it considered the non-stationarity of the data and accounts for the high inter- and intraindividual differences in the temporal course and intensity level of emotions even in the same context. However, an overall mean can still be calculated. Thus, the question should not be if synchrony exists but rather why people differ.

2.2. Manuscript B: Synchrony of Psychophysiological Parameters in Disgust[2]

2.2.1. Purpose of the study

The purpose of Manuscript B was to apply the time-frequency based latent variable approach developed in Manuscript A in order to quantify synchrony of phys-

[2]Deja, M., Muma, M., Hoppe, D. & Kelava, A. (to be revised and resubmit). Synchrony of psychophysiological parameters in disgust. *International Journal of Psychophysiology*, 56 pages.

iological parameters in disgust. With the newly developed quantification method the temporal course of physiological synchrony was analyzed. According to the basic emotion approaches, physiological synchrony should be higher at the beginning of a disgusting picture than at the beginning of a neutral picture (Ekman, 1992). Psychological construction approaches do not necessarily expect a synchronized response during an emotion (Barrett, 2013). Besides the level of synchrony, the interindividual variability and the time course of physiological synchrony were also of interest. The basic emotion approaches would expect similar physiological synchrony responses across several individuals (Ekman, 1992). The psychological construction approaches emphasize interindividual differences in the physiological response to an emotion. Further, the course of physiological synchrony was compared to a continuously recorded subjective rating of disgust.

2.2.2. Method

In total, 42 participants (21 female) watched neutral and disgusting pictures from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2008) in a laboratory setting. The pictures were displayed for 7 s with a black screen (3 s) between them. Physiological parameters such as ECG, respiration, and EDA were measured with the Biopac MP 150 System and the AcqKnowledge 4.2 Software (Biopac Systems Inc, 2011). As in Manuscript A, in a first step, bivariate coherences were calculated. In contrast to Manuscript A, five instead of six bivariate coherences resulted from the calculation. The coherence of the EDA and ECG signals in the pulse frequency range was excluded because these high frequency EDA components are interferences from the ECG signal due to the measurement process rather than physiologically generated signal content (Boucsein, 1988). In a second step, the overall latent physiological synchrony variable was calculated by applying a state-space model. After seeing the neutral and the disgusting picture block, participants retrospectively rated their level of disgust as well as other emotions using the *Differential Affect Scale* (DAS; Merten & Krause, 1993). At the same time, they continuously rated their level of disgust for each picture throughout the course of the experiment using a rating dial. A manipulation check using subjective ratings and physiological features was conducted to check whether disgust had been evoked successfully.

2.2.3. Results and discussion

The manipulation check revealed that the emotion induction led to disgust-specific changes in physiological parameters, decreasing heart rate while increasing electrodermal activity. The subjective rating of disgust was significantly higher than for other negative emotions. In the moment of a picture change, physiological synchrony was significantly higher when the subsequent picture was disgusting instead of neutral. On the one hand, the results showed the existence of an increased physiological synchrony at the beginning of a disgusting picture which supports the basic emotion approach (Ekman, 1992). On the other hand, alternative explanations such as orienting response should be discussed (see section 3). Further, the interindividual variability and temporal development of physiological synchrony was of interest. There were large interindividual differences in the time course and level of physiological synchrony which supports Barrett's (2013) point of view. On average, physiological synchrony already increased when the black screen was presented and reached its maximum in the moment of the picture change, irrespective of the content of the following picture. If the picture after the black screen was neutral, physiological synchrony decreased rapidly and increased again. If the picture after the black screen was disgusting, physiological synchrony declined back to zero within 2 s. The quick decrease of physiological synchrony during the neutral picture could be a form of relief that no disgusting content was displayed. Interestingly, in disgusting pictures synchrony decreased already back to zero, when the subjective rating dial reached its maximum. Hence, subjective experience and physiological synchrony did not reach their maxima at the same time.

2.3. Manuscript C: Person- and Emotion Classification of Fear Using Peripheral Physiological Data[3]

2.3.1. Purpose of the study

In the past, the classification of individuals using physiological parameters has often neglected the high intraindividual variability of physiological parameters in different situations by only using one measurement occasion (e.g., Bailenson et al., 2008; Kolodyazhniy, Kreibig, Gross, Roth, & Wilhelm, 2011; Kreibig, 2010). Therefore, one purpose of the present study was the classification of individuals using data from two measurement occasions to meet the demands of a real-world application to a greater extent. In an analogous manner, the high intraindividual variability of physiological data was neglected in studies regarding the classification of emotions by using one measurement occasion and small sample sizes (4-12), leading to an overestimation of classification results (e.g., C.-Y. Chang, Tsai, Wang, & Chung, 2009; Gouizi, Bereksi-Reguig, & Maaoui, 2011; Jang, Park, Kim, & Sohn, 2012). Hence, the second purpose of the study was the classification of fear using two measurement occasions with a sample size of 36 individuals. In addition, the temporal development of an emotional experience and the course of emotional intensity has just recently been considered in classification research (Walecki, Rudovic, Pavlovic, & Pantic, 2017). Most studies used retrospective subjective data to operationalize emotional intensity (e.g., Rani, Liu, Sarkar, & Vanman, 2006). A third purpose of Manuscript C was to classify different levels of the emotion elicited, and to validate them based on subjective ratings of fear.

2.3.2. Method

In total, 36 participants (20 female) watched neutral and fear-eliciting pictures and film clips on two measurement occasions with a time interval of six weeks between them. Afterwards they were asked to retrospectively rate their level of various emotional states using the *Differential Affect Scale* (DAS; Merten & Krause, 1993). At the same time, they rated their level of fear continuously with a rating dial.

[3]Deja, M., Muma, M., Richter, S., Entringer, T., Hoppe, D. & Kelava, A. (submitted). Person- and emotion classification of fear using peripheral physiological data. *Psychological Test and Assessment Modeling*, 46 pages.

Physiological parameters such as ECG, respiration, and EDA were measured with the Biopac MP 150 System and the AcqKnowledge 4.2 Software (Biopac Systems Inc, 2011). From the ECG features we constructed a feature vector of seven features for classification of individuals. From the ECG, respiratory, and EDA features we constructed a feature vector of nine features for emotion classification according to Gramann and Schandry (2009). Two nonlinear classifiers, namely the k-nearest neighbours (KNN; Cover & Hart, 1967) and the support vector machines (SVM; pioneered by Vapnik, 1998) were applied as classification methods. Further, an extension of the SVM (T. F. Wu, Lin, & Weng, 2004) was used to compare a more continuous classification of fear (probability values instead of 0 and 1) with the subjective intensity ratings of fear.

2.3.3. Results and discussion

The manipulation check showed that the emotion induction of fear was more successful when using film clips rather than pictures as an induction method. First, for the classification of individuals, the random chance of correct classification was 2.78% (since we had 36 individuals). If training and testing data derived from the same measurement occasion, 56.70% (54.53%) of all individuals were correctly classified using the SVM (KNN). There was a high interindividual variability with regard to how good each individual could be classified on the grounds of ECG features. If training and testing data derived from two different measurement occasions, 26.93% (23.16%) of all individuals were correctly identified using the SVM (KNN). Second, as the number of individuals decreased, classification rate increased. Hence, small sample sizes and one measurement occasion have lead to an overestimation of the correct classification of individuals using ECG features in the past. If fear was induced with pictures, the person-dependent classification with a two-fold cross validation correct classification rate reached 66.20% (64.40%) using the SVM (KNN). If fear was induced with film clips, the best correct classification rate was achieved for the SVM with 81.90%. In both scenarios, random chance of classification was 50% (since we only had fear or no-fear as a possible class). The results showed that the emotion induction method and the applied classification method have an influence on the correct classification rate. Third, the two amplitudes of the continuous subjective rating of fear were also visible in peaks of the classification probabilities. Hence, the increase of the probability to belong to the fear class increased with the subjective intensity level of fear. The results of

Manuscript C showed that daily variations across several measurement occasions deserve further study in classification research. The SVM outperformed the KNN. Emotion induction using film clips was more successful than using pictures. It is doubtful whether the probability ratings of the SVM extension may be interpreted as a classification of (fear) intensity levels. The corresponding course of the subjective fear indicator, conveyed by the continuous dial adjustments, however, lends support to that interpretation.

3. General Discussion and Conclusions

This doctoral thesis investigated, on the one hand, the synchrony of physiological response parameters during an emotion (response synchrony) and, on the other hand, the discrimination of an emotion and a neutral state by using physiological parameters (response patterning). For this purpose, three studies were conducted. In total, 136 participants took part in the experiments and three different emotional states, namely amusement, disgust, and fear were elicited using pictures and film clips. This section presents a brief summary of the results. Following, theoretical implications and the resulting consequences for practical application are discussed. Finally, limitations and future directions reported in this thesis are summarized.

3.1. Summary of Results

The goal of Manuscript A was the development of a new method to quantify the synchrony of multivariate response parameters. Therefore, a new time-frequency based latent variable approach was presented. The new approach considered the non-stationarity of the underlying multivariate physiological data and analyzed them on an interindividual as well as on an intraindividual level. Further, the new method was compared with a previously published method of Hsieh et al. (2011). Unlike in the past, the new approach was capable of mapping the individual temporal course of physiological synchrony during the experience of amusement. Thus, interindividual differences in the amount and the temporal course of physiological synchrony were detected. For some individuals, physiological synchrony was close to 1 during the funny film clips. By contrast, the method of Hsieh et al. (2011) only provided probabilities for response synchrony that are not person-specific and do not consider the non-stationarity of the data.

Manuscript B is directly related to Manuscript A. Here, the newly developed quantification method was applied to measure synchrony of physiological response parameters in disgust. The results showed that at the beginning of a disgusting picture, physiological synchrony was at a maximum level and significantly higher than physiological synchrony at the beginning of a neutral picture. However, physiological synchrony in both cases started to increase before the respective picture was presented and thus, can confirm results of previous studies by Bulteel et al. (2014). Obviously, the anticipation of an emotional stimulus (in both cases participants did not know the content of the upcoming picture) already leads to a synchronized response of physiological parameters as a form of preparedness of what might come. Upon closer examination high interindividual variability was found which means not everyone reacted with a physiological synchronized response during disgust. However, physiological synchrony in disgust was detected for a majority of the participants. Further, the time course of physiological synchrony was compared with an online subjective rating of disgust. Interestingly, the maximum level of subjective disgust was experienced 3-4 s later than the maximum level of physiological synchrony. This finding implies that the conscious awareness of disgust is fully developed when the physiological response is already over. Hence, physiological parameters do not change synchronously with the subjective experience.

Rather than on response synchrony, the focus of Manuscript C was on response patterning, hence the direction and not the simultaneity of change of physiological parameters was important. The aim of Manuscript C was to detect fear in comparison to a neutral emotional state based on physiological parameters. Two instead of one measurement occasion for training and testing data were used to better reflect real-world fluctuations of emotional states. Further, to investigate the temporal stability of physiological parameters, individuals were also classified using ECG features. ECG features of each individual from one measurement occasion were used to classify the same individual on the other measurement occasion. The results showed that the correct classification rate for the classification of individuals decreased if training and testing data did not derive from the same measurement occasion. Hence, physiological data alone can vary as a result of the measurement occasion, and thus the starting point for a change in physiological parameters during an emotion is always dependent on the respective measuring time. This should be taken into account when, in the future, machines should recognize individuals or even emotions. A fact that has often been ignored by psychological classification studies. Further results showed that the correct classification rate was higher if film clips instead of pictures were used as emotion induction method. This may be due to the intensity level of fear. The subjective fear rating was higher for the film clips than for the pictures. Finally, in order to classify the intensity of an emotion, an extension of the SVM classifier has been used. The results are encouraging since similar patterns could be extracted for the subjective fear intensity rating and the (objective) fear level classification probability.

3.2. Theoretical Conclusions and Practical Implications

The presented thesis contributes to an ongoing theoretical debate about whether synchrony of response parameters constitutes a core element of an emotional episode, or not. The new approach presented in this thesis is able to compensate for some disadvantages (e.g., overall synchrony variable, time course, intraindividual level, non-stationarity) of the methods used so far. The application of the new approach does, in small steps, contribute to the clarification of the theoretical debate as was demonstrated in this doctoral thesis. Furthermore, the work shows how important it is to consider individual factors such as the time of the measurement when working with physiological parameters. Instead of averaging across individuals, the thesis strongly suggests to analyze physiological data on an individual basis. In the following, the theoretical conclusions and practical implications of the thesis are discussed.

3.2.1. Basic emotion approaches vs. psychological construction approaches

Both of the theoretical approaches discussed make different assumptions regarding the formation of an emotion as well as the functionality of response synchrony. Basic emotion approaches see the response synchrony as a core element of an emotion. Further, the synchronous response should be similar across different individuals (e.g., Ekman, 1992). Psychological Construction approaches do not see the response synchrony as elementary for the experience of an emotion. Changes in different modalities "construct" an emotion with the help of cognitive processes. So far, the temporal course of different response parameters like physiological change and change of subjective experience could not be analyzed due to the absence of a corresponding method.

The results of Manuscript A and Manuscript B revealed large interindividual difference in physiological synchrony during amusement and disgust. Hence, for some individuals an emotional experience was accompanied with physiological synchrony. For others, physiological synchrony did not occur. The basic emotion approaches per se do not assume such differences in the synchronous response, since it fulfills an evolutionary function (Ekman, 1992). Hence, the interindividual differences of physiological synchrony found in Manuscript A and Manuscript B rather support the assumptions of the psychological construction approaches.

However, in Manuscript B, in addition to the interindividual differences, on average, a significantly higher physiological synchrony could be found in disgusting pictures compared to neutral pictures. Further, the standard deviation of physiological response synchrony was smallest when the physiological synchrony reached its maximum level. This means that at the highest level of physiological synchrony the participants responded in the most similar way. Hence, the temporal course of the physiological synchrony to reach a maximum shortly after the picture change and then to decrease again, seems to have been similar in different individuals. The course of the increase and decrease, however, varied. These findings support the assumption of basic emotion approaches that physiological response synchrony during an emotion exists and similarities in the temporal course can be found.

The temporal course of physiological synchrony and the online subjective rating of disgust revealed that the maximum rating of disgust occurred 3 to 4 s after physiological synchrony reached its maximum. The physiological synchrony response happened within the range of 0-4 s which fulfills the criteria of a quick response in basic emotions (Ekman, 1992). Thus, the onset of the subjective response and the increase in psychophysiological synchrony did not occur simultaneously. The results go along with the view of psychological construction approaches (Barrett, 2013; based on the assumptions of James, 1884) that emotions evolve through the process of making meaning of internal body sensations. The basic emotion approaches rather postulate a simultaneous change of physiological and subjective parameters, since they are both activated through the brain (see also the latent variable model of Coan, 2010). However, the cognitive and behavioral processes which lead a participant to turn the rating dial may take more time than the physiological response. Therefore, the basic emotion approach can not be ruled out, either.

The results of Manuscript C showed that the physiological parameters for the same individual measured on different days varied although the situational parameters were comparable. According to the basic emotion approaches the distinctive physiological pattern discriminates emotional states from each other. However, most of psychological studies only used one measurement occasion. Hence, day-to-day variations of physiological data have not been considered. Compared to previous studies, the correct classification rates reported in this thesis are lower than in other studies (e.g., Kolodyazhniy et al., 2011). This might be due to the two measurement occasions investigated or the less successful emotion induction. However, the rather low classification rates support the assumptions of the psychological construction approaches. The approaches include individual and situational influences on physiological parameters and similar reaction patterns are not necessarily expected (Barrett, 2006a). On the other hand, in the basic emotion approaches, once the fear circuit is activated the response should be similar for different individuals in different situations (Ekman, 1992).

Summing up, some of the results support the basic emotion approaches, namely on average a significant increase of physiological synchrony during disgust. However, the results of this doctoral thesis partly support the assumptions of the psychological construction approaches as well, namely the high interindividual differences. The low classification rates and the time differences between the physiological synchrony and subjective rating dial may be explained by theoretical assumptions of the psychological construction approaches. On the other hand, there are other reasonable explanations such as failed emotion induction or cognitive and motor processes. Hence, the assignment of which approach can be supported is not clear.

3.2.2. Practical implications

The introduced approach to quantify physiological synchrony has also interesting applications in the context of affective computing or the classification of emotions, respectively. According to the basic emotion approaches, parameters of different modalities (e.g., behavior and physiology) should change synchronously during an emotional experience (Ekman, 1992). From this point of view, the classification accuracy should increase if more than one modality is used for the classification of an emotion. Some studies report an increase in classification accuracy if more than one modality is used (e.g., Abdat et al., 2011; Bailenson et al., 2008). D'Mello and Kory (2015), however, reported only modest multimodal effects in their meta-analysis. Their findings stand in line with the psychological construction approaches (Barrett, 2013; Coan, 2010). In future research, the present, newly developed approach might help to determine which modalities change synchronously - and to what extent - and which do not.

Furthermore, the quantification of synchrony could be important in the area of clinical psychology. H. S. Schaefer, Larson, Davidson, and Coan (2014) showed in their study that phobic participants had a greater synchrony between neural activation, physiological parameters, and self-report. However, it is unclear whether synchrony is an indicator of psychopathology or just of the much more intense sensation the patients tend to experience. Further, borderline personality disorder (BPD) is supposed to be associated with a higher emotional reactivity, a higher baseline emotion intensity, and maladaptive emotion regulation (Linehan, 1993). Some empirical studies find differences in the physiological and subjective reaction towards emotion eliciting stimuli between patients with BPD and control groups (Aleknaviciute et al., 2016), while others only report deviation only in sole parameters (Dixon-Gordon, Yiu, & Chapman, 2013), or no differences are reported (Herpertz & Koetting, 2005). Hence, the research of physiological and subjective synchrony during baseline and emotion eliciting conditions could provide valuable information about the emotion processing for patients with BPD. In general, the approach could be applied in future studies to help answer the question if the synchrony of response parameters during emotions differs in patients with a psychiatric diagnosis.

Further, Sze et al. (2010) showed in their study that synchrony between cardiac responses and subjective ratings during an emotion was higher for meditators and dancers than for control subjects. They conclude that participants with a higher body awareness have a more synchronous response in emotion. If response synchrony is a trainable and changeable variable, it could be important for biofeedback applications (Thompson & Thompson L., 2003).

3.3. Limitations and Future Perspectives

In the following section the limitations in terms of emotion elicitation and methodological issues are discussed. Furthermore, future perspectives are presented on the basis of the results of this doctoral thesis.

3.3.1. Emotion elicitation

In this doctoral thesis, amusement, disgust, and fear were induced in a laboratory setting by using pictures and film-clips. In an artificial setting like the laboratory, emotion might not be as naturalistic and intense as the daily emotional experience (Wilhelm & Grossman, 2010). The emotion elicitation of fear with pictures revealed very low subjective intensity ratings. Whether the emotion was induced with sufficient intensity, is doubtful and might be reflected in the low correct classification rates of fear when using pictures. On the other hand, the induction of disgust with pictures led to high subjective intensity ratings. The use of other physiological sensors, which can be more seamlessly connected to the participants, such as wearable heart rate sensing devices might make it possible, in future studies, to induce emotions in a more realistic setting (Schäck, Sledz, Muma, & Zoubir, 2015). However, a more realistic setting for emotion induction often goes along with an increase of body movements and outside noise, which reduces the validity of the experiment. Further, physiological sensors are sensitive to body movements, which could make the attribution to an emotional reaction more difficult.

In each of the three presented studies in this doctoral thesis, only one emotion was induced at a time. Further, in Manuscript B and Manuscript C the different intensity levels within the respective emotion were important. One might argue that in each experiment different emotional states should haven been induced in order to see whether, for example, fear or anger was measured. On the other hand, well validated and often applied material was used for the emotion induction in each study. Furthermore, the discrimination between arousal and, for example, fear was not made in this thesis. Arousal and intensity rating of an emotional state is seen independently (e.g., Reisenzein, 1994) because one can have a high intensity rating of boredom but a low arousal (Lyusin & Ovsyannikova, 2016). No supplemental material for the detection of valence and arousal was provided in the here presented studies. In future studies, different emotional states might be induced as well as complementary material for the assessment of valence and arousal. In this way, physiological synchrony can be compared for different emotional experiences and for different levels of valence and arousal. Another relevant factor should be mentioned here. The majority of studies that classified emotions based on physiological parameters induced at least one of the basic emotions (fear, anger, disgust, sadness, happiness). However, in his meta-analysis D'Mello (2013) points out that emotional states such as boredom, frustration, or engagement do not fulfill all criteria of the basic emotions but are often more important for the user devices like computer games or learning tutorials than the basic emotions themselves. Hence, in the future, application-oriented emotions should be induced.

In the study of physiological synchrony, large interindividual differences were observed which is in accordance with previous empirical results (Bulteel et al., 2014; Hsieh et al., 2011; Mauss et al., 2005). Possible reasons for the interindividual differences are dispositional factors (e.g., interoceptive sensitivity; see also Barrett, Quigley, Bliss-Moreau, & Aronson, 2004; Herbert, Pollatos, & Schandry, 2007) and situational factors (e.g., social acceptance of expression an emotion; see also Cacioppo et al., 1992). However future studies are necessary to investigate the reasons of the interindividual differences. Further, sex differences might account for the variance as well (Poláčková Šolcová & Lačev, 2017). Previous studies regarding synchrony mostly only had female participants (Bulteel et al., 2014; Butler et al., 2014; Gentsch et al., 2014; Mauss et al., 2005; H. S. Schaefer et al., 2014). Some results from gender research show that women have a higher emotional expressibility (Chaplin, 2014; Kring & Gordon, 1998) and show a higher increased skin conductance level in disgust (Rohrmann & Hopp, 2008). Gender differences in physiological and subjective synchrony in disgust or amusement was not the research focus of this doctoral thesis, but should be considered in future studies.

3.3.2. Methodological issues

The studies presented in this cumulus also have methodological limitations that will be discussed in the following section.

In general, the two studies reported concerning the research of synchrony limited the investigation to the synchrony of physiological parameters. Hence, the results can not be transferred to the synchrony between different response modalities, for example, the behavioral and physiological response modality. However, with the approach presented in this thesis it will be possible to examine synchrony between response modalities on an inter- and intraindividual level. Further, the three physiological parameters (heart, electrodermal activity, and respiration) were selected, as they were repeatedly used in research to discriminate between different emotions (Kreibig, 2010). Of course, there are other physiological parameters that can be included as indicator variables for a latent variable model, such as change in electromyographic activity or finger temperature. Moreover, the connection between brain activity and physiological changes would also be of interest. Using the new approach, indicators of e.g. an electroencephalography can be inserted into the model or used exclusively.

The presented quantification approach in Manuscript A accounts for the non-stationarity of the physiological data. However, synchrony of physiological parameters might be more complex and nonlinear (Lewis, 2005). The applied synchrony measure is able to reveal linear interactions between the signals. The approach will be extended to capture nonlinear interactions between the signals, even in presence of motion artifacts, which will allow for assessing physiological synchrony in a more naturalistic setting (see Schäck, Muma, Feng, Guan, & Zoubir, 2017). Further, the approach can not include signals of constant amplitude, such as the data of the rating dial (similar to calculating the correlation between numbers that change with a high frequency and numbers that change with a low frequency). Participants continuously rated their subjective experience during the emotional stimuli. However, they did not constantly change their subjective rating, hence, values of the rating dial remain constant for a certain time period. For such time intervals, the proposed approach can not be used for the quantification of synchrony of physiological and subjective data. For the time being, this problem remains unsolved.

In retrospective, the study designs of Manuscript A and B can be improved for future studies. In Manuscript A two funny film clips were shown. In this case, it is very difficult to define when an emotion starts, when it ends and at what point in time the participants found the respective film clip funny. A rating dial as in Manuscript B would have helped. Further, it is important to include neutral stimulus material. Through the comparison of neutral and emotion eliciting material, changes in physiological parameters can be attributed to the content and not to a stimulus itself. In Manuscript B, pictures of neutral content were shown in a baseline block and pictures of disgusting and neutral content were shown in the experimental block. From 15 disgust pictures in total, the five most disgusting pictures were chosen to ensure that emotions were evoked on a high intensity level. In future, the number of pictures should be reduced to two to three neutral-disgusting picture sequences and the disgusting picture block should also include neutral-neutral picture sequences. In this form, it can be analyzed whether the same response pattern can be found for neutral-neutral picture sequences when the expectancy of a disgust picture might be higher. Moreover, following studies should include pictures with a positive valence to see whether the early increase of physiological synchrony can also be observed here (see also Bulteel et al., 2014). In this scenario, anticipatory anxiety should be ruled out as a possible explanation. Further, neutral and disgusting pictures were shown for 7 s with a 3 s black screen in between. The chosen time frames were similar to other experimental designs (e.g., Lang, Greenwald, Bradley, & Hamm, 1993; Stark, Walter, Schienle, & Vaitl, 2005). However, it may be possible that the total time of 10 s between the pictures was not long enough to ensure that the arousal from the previous picture did not interfere with the arousal of the following picture. Hence, in the future, as in the experimental design of Manuscript C, pictures should have a longer exposure time with a longer black screen between them. At last, it was not controlled for possible avoidance behavior of the participants in Manuscript B and Manuscript C (e.g., turning away, closing eyes). In future studies, avoidance behavior should be controlled for, to prevent emotion regulatory effects on the physiological data.

For the classification of fear, two well-established classifiers from machine learning (kNN and SVM) were used in this doctoral thesis. Both classifiers are still being used (e.g., Jatupaiboon, Pan-Ngum, & Israsena, 2015; Li et al., 2016). However, recently more sophisticated methods have been applied, such as deep neural networks (e.g., Al-Nafjan, Hosny, Al-Wabil, & Al-Ohali, 2017). Further, feature selection algorithms (Swangnetr & Kaber, 2013; Wei & Jia, 2016) were not applied to the data, since emotion relevant features according to Gramann and Schandry (2009) were a priori selected. Since the applied features were not the main research question, individual feature performance in classification and the rank of features were not presented. In future studies the respective feature performance should be presented as well. In general, a standardization of physiological features used for classification analysis should be the focus of future studies. Moreover, a person-dependent emotion classification approach was used. This means that training and testing data derived from the same person. Of course, for applications such as emotion detection in vehicles of a car sharing provider, a person-independent classification is necessary. Here, an unknown person would be classified with the help of a large training data base (Chueh et al., 2012). Because of the large interindividual differences in physiological parameters this is a difficult task but the problem should be addressed in the future. Further, an extension for the SVM classifier was used (T. F. Wu et al., 2004) to yield probability values for the respective class instead of nominal zero/one results. Therefore, film clips were divided into 15 s segments. To ensure a classification of emotions that is closer to real-time classification, the time segments should be reduced.

Summing up, the main contributions of this thesis were the presentation and application of a new approach to quantify synchrony of response parameters in emotion. Despite theoretical controversies, physiological synchrony in disgust and amusement were shown. The classification rates presented in this thesis were rather low but emphasized the difficult undertaking of classifying emotional states based on physiological data on two different measurment occasions.

4. Manuscript A: A new Approach for the Quantification of Synchrony of Multivariate Non-Stationary Psychophysiological Variables During Emotion Eliciting Stimuli

Abstract

Emotion eliciting situations are accompanied by changes of multiple variables associated with subjective, physiological and behavioral responses. The quantification of the overall simultaneous synchrony of psychophysiological reactions plays a major role in emotion theories and has received increased attention in recent years. From a psychometric perspective, the reactions represent multivariate non-stationary intraindividual time series. In this paper, a new time-frequency based latent variable approach for the quantification of the synchrony of the responses is presented. The approach is applied to empirical data, collected during an emotion eliciting situation. The results are compared with a complementary interindividual approach of Hsieh et al. (2011). Finally, the proposed approach is discussed in the context of emotion theories, and possible future applications and limitations are provided.

4.1. Introduction

Researchers agree that emotion eliciting situations are accompanied by changes of multiple variables associated with subjective, physiological and behavioural responses. The reasons for a possible coupling of the response variables is a topic of ongoing discussion in various emotion theories. There is no uniform terminology to describe the simultaneous changes of the response variables. The term 'coherence' is frequently used (e.g., Dan-Glauser & Gross, 2013; Herring, Burleson, Roberts, & Devine, 2011; Hsieh et al., 2011; Mauss et al., 2005; Reisenzein, 2000; Rosenberg & Ekman, 1994; Sze et al., 2010) to describe the simultaneity of changes in the response variables. Further terms that are used in the research community to describe the interrelation of the responses are 'synchronization', 'organization of response systems', or 'concordance'. Throughout this paper, the terms synchronization and synchrony are interchangeably used to describe the simultaneous changes of response variables. This article introduces a new approach for the quantification of the synchrony of the response variables that is able to account for non-stationarity. A signal is non-stationary if its mean and covariance function are time-varying (Brillinger, 2001). First, different assumptions regarding the functionality of the synchrony concept in different emotion theories are explained. Subsequently, the new approach, which basically consists of two steps, is introduced. In a first step, time-frequency based bivariate coherence measures are derived (Muma et al., 2010). In a second step, these measures are used in an state space modeling approach to obtain an overall synchrony measure of the simultaneous activation of psychophysiological responses. The approach is then applied to empirical data collected during an emotion eliciting situation and compared with a complementary approach of Hsieh et al. (2011). Finally, the proposed approach is discussed in the context of emotion theories, and possible future applications and limitations are provided.

On the functionality of synchrony of responses in emotion theories

Emotion theories make different assumptions regarding the functionality of a synchrony of response variables: Basic emotion theories state that different emotions have distinct and coordinated patterns of physiological responses. According to basic emotion theories, the specific psychophysiological response variables are activated simultaneously during an emotional experience but are less associated with each other during rest (Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962). A synchronized response is desirable as it prepares the body for an adequate reaction to a stimulus and leads to an appropriate reaction to environmental demands (Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962). Different emotional responses are organized by central mechanisms in the brain, such as the amygdala (e.g., LeDoux, 2000; Murphy et al., 2003; Whalen et al., 1998), orbitofrontal cortex (e.g., Goodkind et al., 2012; Hornak, 2003; Murphy et al., 2003), the insular (e.g., Damasio et al., 2000; MacLean, 1990; Murphy et al., 2003), and other brain regions (e.g., Ekman, 1992; Izard, 2011; Panksepp, 2008). The synchronized specific responses in time and intensity have been interpreted as evidence for the existence of a causal mechanism (e.g., Kettunen, Ravaja, Näätänen, Keskivaara, & Keltikangas-Järvinen, 1998; Levenson, 2003). Further, emotions can also activate so called 'affect programs' that include behavioral and physiological changes, that might be similar for different individuals (Ekman & Cordaro, 2011; Tomkins, 1962). For example, fear, anger, or amusement may be accompanied by a specific reaction pattern in terms of subjective experience, physiology, and behavior (Ekman, 1992; Levenson, 2011; Murphy et al., 2003; Panksepp & Watt, 2011).

A different viewpoint is taken by so-called dimensional approaches that do not classify the emotional experience in distinctive categories, such as anger or fear. Instead, these approaches discriminate between different emotional states by introducing a two-dimensional space that is spanned by valence (pleasure/displeasure) and arousal (activated/deactivated) as sufficient means to discriminate between different emotional states (Barrett, 2006a; Barrett & Russell, 1998; Russell, 1980, 2003). Higher dimensional spaces have been proposed by, e.g., Bradley and Lang (1994) or Fontaine, Scherer, Roesch, and Ellsworth (2007). The assumption that an event causes an emotion and the emotion causes a synchronized, specific change in cognition, behavior, and physiological reaction is criticized by Russell (2003). Instead, in the conceptual framework of the core affect, the hypothesis is made that that emotions do not have a common cause. From this perspective, a synchrony between the responses is not required (see also Barrett, 2006a; Russell, 2009). On grounds of genetic differences, personal experiences (e.g., Lykken & Tellegen, 1996), responsiveness to stimuli, attributions, and other factors, individuals can respond differently to the same emotion eliciting situation (Russell, 2003). Hormonal changes, endocrine dysfunction, illness, satiety and diurnal rhythms can internally influence the emotional response (Russell, 2003), which results in very specific patterns of psychophysiological variables on the intra- and interindividual level.

A third body of research relies on the concept of appraisal. Appraisal indicates an evaluation of the situation regarding its personal significance (Lazarus, 1991) and therefore, in the same situation there can be a great variation between the emotional state of individuals (Kuppens, Stouten, & Mesquita, 2009; Scherer, Dan, & Flykt, 2006). Still, a synchronous response during an emotional episode is expected (Scherer, 2009). Some of the appraisal theories view the appraisal of a situation, and not the event itself (as projected by the basic emotion theory; Russell, 2003), as a causal mechanism responsible for the elicitation of an emotion (e.g., LeDoux, 1989; Roseman et al., 1990; Schachter & Singer, 1962; Scherer, 1993). Lewis (2005), on the other hand, assumes a recursive, complex relationship between subsystems of the nervous system and hence, not a linear relationship of one causal mechanism followed by a cascade of responses. Grandjean et al. (2008) consider various feedback loops between the synchronized response of the peripheral-, motivational-, monitor-, cognitive-, and motor-system to be responsible for the conscious awareness of an emotion (e.g., Fries, 2005; Scherer & Ekman, 1984; Scherer, 1987, 2009). The synchrony of multiple response variables, however, is considered to be necessary for a conscious emotional experience (Grandjean et al., 2008). Also, several appraisal approaches support discrete categories of emotion (e.g., Oatley & Johnson-laird, 1987; Roseman et al., 1990) while others approve the dimensional perspective (e.g., Kuppens, Champagne, & Tuerlinckx, 2012; Scherer et al., 2006; L. Wu, Winkler, Andreatta, Hajcak, & Pauli, 2012). The above approaches are theoretical and the question remains how to quantify the synchrony of the response variables. The specific response variables constitute multivariate time series with time-varying distributions. Additionally, Dan-Glauser and Gross (2013) recently stated that defining a measure of synchrony is a challenging and timely topic in emotion theory. In the following section, the results concerning the synchrony of peripheral physiological response variables during emotion eliciting situations are reviewed.

The synchrony of peripheral physiological response variables during emotion eliciting situations: A brief review of results

In early psychological research of emotion and stress, the synchrony of peripheral physiological measures was a major focus of the analyses (Lacey & Lacey, 1958; Lazarus et al., 1963; Nesse et al., 1985; Wenger, 1942). For example, the autonomic response system (ANS) plays an important role during stress (e.g., Bibbey, Carroll, Roseboom, Phillips, & de Rooij, 2013; Carroll, Lovallo, & Phillips, 2009), posttraumatic stress disorder (e.g., Ehlers et al., 2010; Zucker, Samuelson, Muench, Greenberg, & Gevirtz, 2009) or anxiety-disorders like panic-attacks (e.g., Meuret et al., 2011; Roth, Telch, Barr Taylor, & Stewart Agras, 1988). Information on the temporal interdependences of peripheral physiological measures can enhance the understanding of the underlying functioning of the ANS (Kettunen et al., 1998; McAssey, Helm, Hsieh, Sbarra, & Ferrer, 2013) and can provide important information about the psychophysiological processes during an emotional episode (McAssey et al., 2013). However, it is unclear under which conditions, and to what exact quantitative extent, a synchronous physiological reaction occurs (Hsieh et al., 2011; McAssey et al., 2013). Applying intraindividual time series models, Kettunen et al. (1998) reported that the synchronization between electrodermal activity and heart rate within an individual is associated with a higher level of arousal and behavioral activity (see also Lazarus et al., 1963). In their stochastic network configuration approach, Hsieh et al. (2011) found three clusters within an overall system that is formed by 15 psychophysiological signals: one behavioral cluster and two physiological clusters (blood pressure and cardiovascular parameters). They reported a higher synchronization between the different clusters as the intensity of an emotional stimuli was measured from a neutral condition, suggesting that there is a higher association during an emotional experience episode. A higher synchrony was also reported within each cluster. Based on brain activity analysis, Costa, Rognoni, and Galati (2006) found by using a synchronization index, a higher synchrony of various EEG channels during emotional film stimuli than during neutral film clips. Their results indicate a higher information exchange during emotional responses (for similar results see also e.g., Miskovic & Schmidt, 2010).

Numerous studies on psychophysiological correlates of emotional stimuli have been undertaken. However, reactions from emotion eliciting stimuli are not universal on the inter- and intraindividual levels. Individuals vary in the intensity and duration of an emotional episode in terms of subjective experience, physiological and behavioral reactions (e.g., Grandjean et al., 2008; Kuppens et al., 2009). Thus physiological response patterns during emotion tend to differ between individuals (e.g., Kristjansson, Kircher, & Webb, 2007; Marwitz & Stemmler, 1998). Kristjansson et al. (2007) applied a two level growth curve model, in which the first level explains the variance within participants and the second level the variance between participants. Marwitz and Stemmler (1998) used correlations and ANOVA to analyze individual response specificity. The physiological responses were found to be influenced by the individual appraisal (see Scherer, 2009, for an overview), emotion regulation (e.g., Dan-Glauser & Gross, 2011; Gross & Levenson, 1993, 1997), and context specific attributes (see Cacioppo et al., 1992, for an overview). Dan-Glauser and Gross (2011), Gross and Levenson (1993), and Gross and Levenson (1997) used ANOVAs to detect the effect of emotion regulation on physiological data. Dan-Glauser and Gross (2013) showed in their study, by applying cross-correlations, that synchrony within the physiological channel decreased, if participants were instructed to suppress their emotions. Lacey and Lacey (1958), on the other hand, showed that the physiological response can also vary within an individual (see also Cacioppo et al., 1992). As described in the previous section, basic emotion theory, as well as some of the appraisal approaches, suggest a higher intraindividual synchrony during emotion. Nevertheless, some responses, e.g., respiratory and cardiovascular measures (respiratory sinus arrhythmia; RSA) are also synchronized in order to assist biological functions within our body (Ben-Tal, Shamailov, & Paton, 2012; Garcia, Koschnitzky, Dashevskiy, & Ramirez, 2013; Yasuma, 2004).

In contrast to the large number of studies that examine the correlates of emotional response systems to emotional stimuli (Bonanno & Keltner, 2004; M. G. Calvo & Miguel-Tobal, 1998; Dan-Glauser & Gross, 2013; Hsieh et al., 2011; Mauss et al., 2005; Reisenzein, 2000; Rosenberg & Ekman, 1994; Sze et al., 2010), there are only a few empirical studies that provide a quantitative measure of synchrony (Kettunen et al., 1998; Lacey & Lacey, 1958; Lazarus et al., 1963; Nesse et al., 1985; Wenger, 1942). There may be several reasons for this: Firstly, the analysis of physiological response variables is difficult, since they require multivariate, nonlinear, and non-stationary analysis methods (Zong & Chetouani, 2009). Non-stationarity arises when the joint probability density function (pdf) of the response variables changes over time (see next section for more details). Most of the approaches applied so far e.g., cross-correlation, implicitly rely on the stationarity of the physiological signals, and such an assumption is not fulfilled in practice (Muma et al., 2010).

Secondly, specifying a model that quantifies a time-varying synchrony of multiple response variables is not trivial (Dan-Glauser & Gross, 2013).

Thirdly, not only the psychophysiological responses, but also the emotions, impose additional challenges and the demand for sophisticated analytical methods. Emotions have often been treated as static phenomena, similar for different individuals, and have been analyzed by using nomothetic approaches, neglecting the intraindividual variability and the dynamic process of an emotional experience (Kuppens et al., 2009). interindividual analysis of the mean values of physiological responses overlooks the possible synchronous response within an individual over time. Therefore, an intraindividual analysis is more appropriate for the analysis of synchrony (Hsieh et al., 2011; Lazarus et al., 1963; Mauss et al., 2005). In the following section, a new approach, which takes the non-stationarity of the data into account, is introduced for analyzing multivariate synchrony of peripheral physiological measures in an individual during an emotional event.

4.2. A New Approach for the Quantification of Synchrony of Multivariate Psychophysiological Signals

In this section, we introduce a new time-frequency based latent variable approach for the quantification of the synchrony of peripheral physiological responses, such as the activity of the heart, respiration, and the electrodermal activity level. The concepts of spectral bivariate coherence and time-frequency bivariate coherence from a signal processing perspective are first discussed. These concepts provide the basis for the quantification of synchrony of non-stationary psychophysiological response variables. Bivariate coherences are used as indicators of a latent state-space model, which quantifies one latent synchronized variable[4].

[4]In this paper a distinction between spectral coherence and synchrony is made. Spectral coherence is a well defined bivariate measure in the signal processing literature (Brillinger, 2001; Marple, 1987). The term 'synchrony' defines an underlying latent state of an overall system-wise synchrony.

Spectral bivariate coherence

A frequently used function to examine the linear relation at frequency f between two signals $x(t)$ and $y(t)$, which are a function of time t, is the spectral coherence $C_{XY}(f)$ (Brillinger, 2001; Marple, 1987), which is defined according to

$$C_{XY}(f) = \frac{S_{XY}(f)}{\sqrt{S_{XX}(f)S_{YY}(f)}}. \tag{4.1}$$

Here, $S_{XY}(f)$, $S_{XX}(f)$, and $S_{YY}(f)$ denote the cross-spectrum and the auto-spectra of $x(t)$ and $y(t)$, respectively. The signals $x(t)$ and $y(t)$ can, e.g., be two different psychophysiological time series measurements. In practice, spectra can be estimated, e.g. based on the periodogram, which is computationally efficient, since it uses the fast Fourier transform (FFT; Brigham, 2002). The raw periodogram is not a consistent spectral estimate, since its mean-squared error does not decrease to zero as the number of samples used in the computation increases to infinity (Marple, 1987). Consistent spectral estimates can be obtained, e.g. by Welch's and Bartlett's methods, by approaches that: (i) split the original measurement into K segments, for which the periodogram is computed and (ii) obtain the overall spectral estimate by averaging over the K periodograms of the segments. Alternatively, consistency can be achieved by the Blackman-Tukey estimator which performs a smoothing of the periodogram. This is achieved by a convolution of the periodogram with a spectral window to reduce the variance. In the time-domain, this operation corresponds to a multiplication of the sample covariance sequence with a lag window of smaller size than the data size (Stoica & Moses, 2005).

$C_{XY}(f)$ displays the consistency-of-phase-relationship between $x(t)$ and $y(t)$ and provides a frequency selective measure of the phase coupling between the two signals (Brillinger, 2001).

The values of the magnitude squared coherence $|C_{XY}(f)|^2$ will always satisfy the relationship $0 \leq |C_{XY}(f)|^2 \leq 1$. Since $C_{XY}(f)$ is normalized by the product of the auto-spectra (see Equation (4.1)), it is independent, e.g., of different amplitudes of $x(t)$ and $y(t)$. If $x(t)$ and $y(t)$ are completely uncorrelated, their coherence will be zero. If, on the other hand, $|C_{XY}(f)|^2 = 1$ then $y(t)$ can be fully predicted from $x(t)$ by a linear and time-invariant system (Marple, 1987). The case $|C_{XY}(f)|^2 < 1$, as described in Bendat and Piersol (1990), may be either due to: (i) noise entering the measurements, (ii) a non-linear functional relationship between the signals, or (iii) further inputs in addition to $x(t)$ contributing to the output. It should be emphasized, however, that the above mentioned statements about $C_{XY}(f)$ are only valid for second order stationary signals (Bendat & Piersol, 1990; Brillinger, 2001; Marple, 1987). Second order stationarity implies a constant mean and variance, as well as an autocovariance function that does not depend on t. To overcome the limitations that arise from the assumption of stationary signals, coherence measures for non-stationary signals have been developed. These are based on the wavelet transform (Grinsted, Moore, & Jevrejeva, 2004) and time-frequency distribution based methods (Matz & Hlawatsch, 2000; Muma et al., 2010; Orini, Bailón, Mainardi, Laguna, & Flandrin, 2012; White & Boashash, 1990). The following example illustrates some limitations of classical spectral coherence analysis and provides motivation for the use of time-frequency representations. Figure 4.1 displays the time, frequency, and joint time-frequency representations of two non-stationary signals $x(t)$ and $y(t)$.

- $x(t)$ consists of a superposition of a linear frequency modulated signal (a signal that consists of a single frequency component whose instantaneous frequency changes linearly over time; starting at 1 Hz for $t = 0$ and reaching 2 Hz at $t = 20$ seconds) and two sinusoidal signals with frequencies of 3 Hz and 4 Hz.

- $y(t)$ consists of a superposition of a linear frequency modulated signal (an instantaneous frequency starting at 1 Hz for $t = 0$ and reaching 2 Hz at $t = 20$ seconds) and a sinusoidal signal with a frequency of 4 Hz and of 10 seconds duration.

The time-frequency representation, like a sheet of music, displays the frequency content of a signal evolving over time. This wealth of information is lost when only the frequency domain is considered, since in the spectrum estimation process, averaging over time is performed. The loss of information is passed on to the spectral coherence measure $C_{XY}(f)$. Figure 4.2 (first panel) displays the spectral coherence estimate for the above example. $C_{XY}(f)$ only indicates that there is a high synchrony between $x(t)$ and $y(t)$ in the frequency region between 1 and 2 Hz, but the evolution/progression over time is averaged out. Furthermore, $C_{XY}(f)$ takes a value of about 0.5 at a frequency of 4 Hz, suggesting a moderate coupling between the signals. This average value neglects the fact that the coherence is nearly equal to one for half of the time and zero for half of the time. These examples illustrate, if the signals are non-stationary, i.e. their spectra evolve over time, that it is important to take the time variation information into account when establishing measures of coherence.

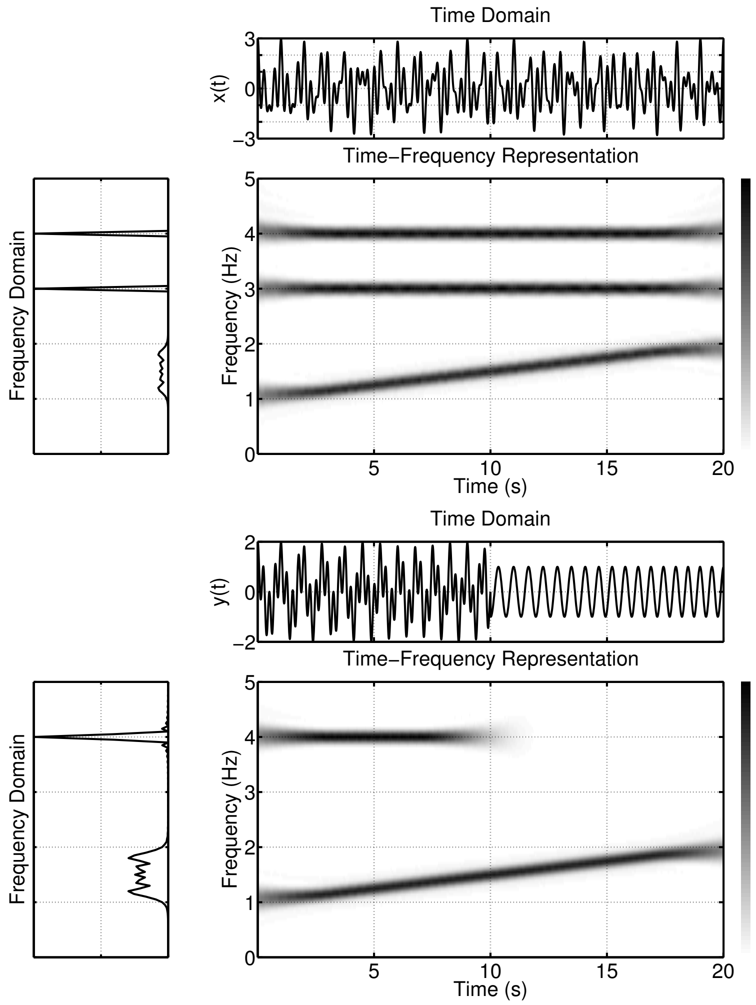


Figure 4.1. Example of two non-stationary signals in time, frequency and time-frequency domains.

Time-frequency bivariate coherence

The notion of time-frequency (TF) coherence was first defined by White and Boashash (1990) and then extended by Matz and Hlawatsch (2000). The definition is

$$C_{XY}(t,f) = \frac{S_{XY}(t,f)}{\sqrt{S_{XX}(t,f)S_{YY}(t,f)}}, \tag{4.2}$$

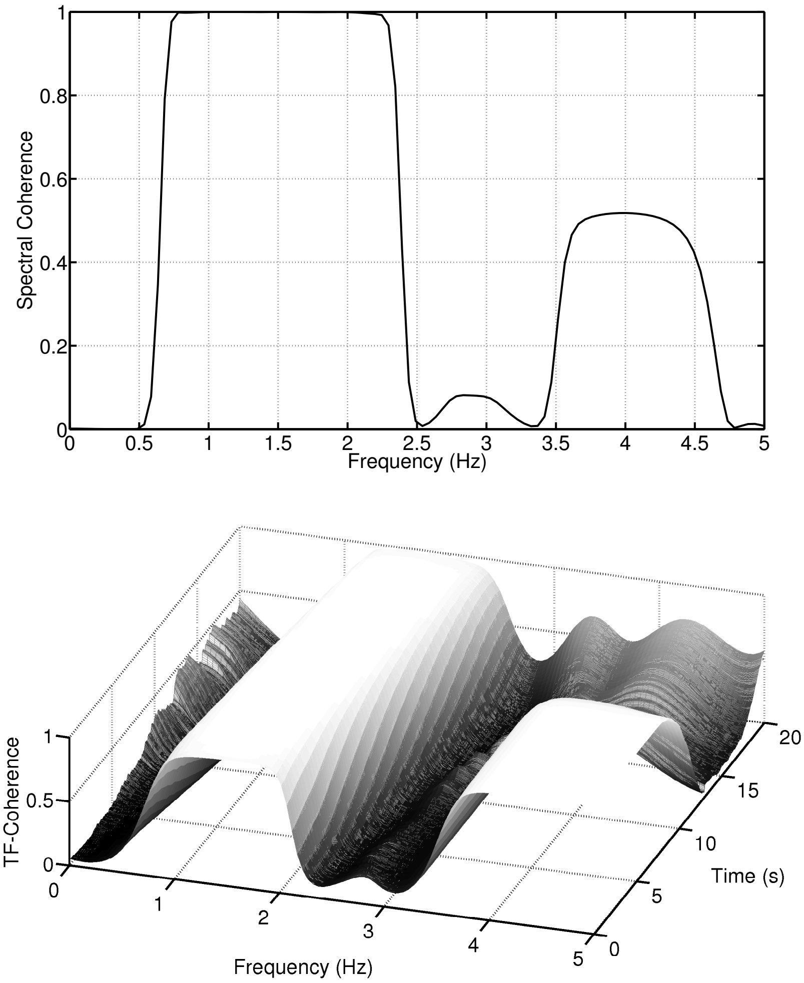


Figure 4.2. Spectral coherence and time-frequency coherence of the simulated non-stationary signals.

where $S_{XY}(t,f)$, $S_{XX}(t,f)$ and $S_{YY}(t,f)$ are the cross and auto time-frequency distributions of $x(t)$ and $y(t)$, respectively. There exist some conditions, which are stated in Matz and Hlawatsch (2000), that are necessary for $C_{XY}(t,f)$ to be well defined and produces meaningful results. In particular, these conditions guarantee that $0 \leq |C_{XY}(t,f)|^2 \leq 1$, where $|C_{XY}(t,f)|^2 = 0$ for uncorrelated signals and $|C_{XY}(t,f)|^2 = 1$ if $x(t)$ and $y(t)$ are related via a linear time-invariant filter of sufficiently short length compared to the stationarity width (for details see White & Boashash, 1990).

The time-frequency distribution used in this paper, is the spectrogram which satisfies the conditions for $C_{XY}(t,f)$ to be well defined and has been suggested for use in TF coherence estimation by White and Boashash (1990). The spectrogram is based on the short-term Fourier transform (STFT), and is simply a sequence of FFTs of windowed data segments, where the windows overlap in time. The spectrogram is defined as the magnitude square of the elements of the STFT. The spectrogram yields a time-frequency plot which contains columns of spectral estimates for a specific moment in time. In addition to the choice of the nature of the time-frequency distributions that underpin $S_{XX}(t,f)$, $S_{YY}(t,f)$ and $S_{XY}(t,f)$, it is also necessary to perform a smoothing operation on the distributions, Matz and Hlawatsch (2000). In this paper, as in Muma et al. (2010), the smoothing is performed by a 2 dimensional filtering with a Gaussian kernel as shown in Figure 4.3. The parameters of the Gaussian kernel define the time and frequency resolution capability of $C_{XY}(t,f)$ and have been chosen such that both sharp changes in time (time resolution) and individual frequency regions of interest (frequency resolution) can be resolved. Figure 4.2 (second panel) illustrates the resolution capabilities based on the signals displayed in Figure 4.1. It can be seen that $C_{XY}(t,f)$ is able to display the time-varying coherence of the two non-stationary signals. As discussed earlier, this is not possible with the spectral coherence $C_{XY}(f)$, alone. Similarly, time domain methods, that rely on stationarity, e.g. correlation coefficients, cannot adequately describe the time-varying relationship between these two exemplary non-stationary signals. In particular, computing the correlation coefficient relies on a non time-varying correlation function (i.e., stationarity).

To illustrate the usefulness of time-frequency based methods for analyzing psychophysiological data, three synchronously measured signals have been considered: An electrocardiogram (ECG) signal, a respiratory signal, and a signal measuring the electrodermal activity level (EDA). The results are shown in Figure 4.4. Vi-

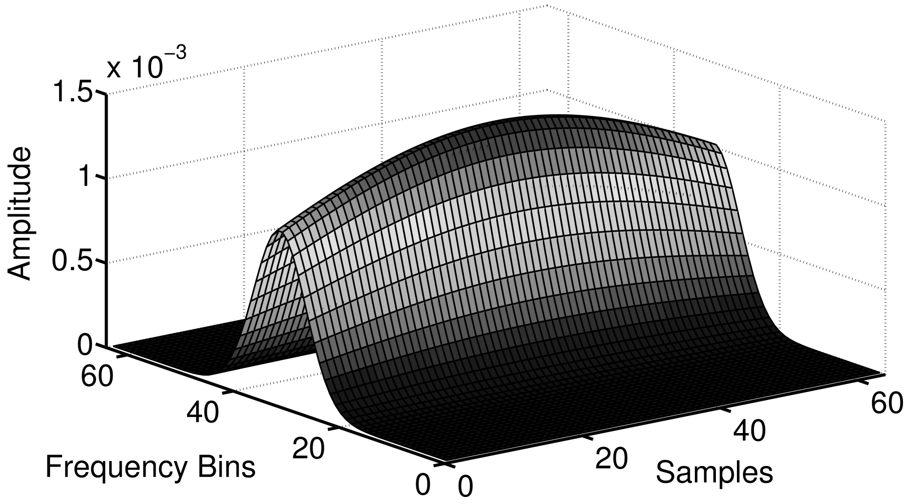


Figure 4.3. The implementation of the Gaussian smoothing kernel. Smoothing of the time-frequency distributions is necessary for $C_{XY}(t, f)$ to be well defined and produce meaningful results.

sual inspection quickly reveals that the spectra change over time, indicating the non-stationarity of the signals. A more formal testing for stationarity of some physiological signals, among them ECG signals and respiratory signals, has been performed in Muma et al. (2010) using a frequency domain test, suggested by Brcich and Iskander (2006). The test is able to determine if stationarity exists at a given frequency region of interest. The test, in general, rejected the hypothesis of stationarity for ECG and respiratory signals.

As can be seen in Figure 4.4, the power of the signals is not evenly spread out in the time-frequency plane, but instead is concentrated in delineated regions. If only one such region exists, the signal is referred to as 'mono-component', while if multiple delineated regions exist, the signal is referred to as 'multi-component'. For example, a linear frequency modulated signal is a mono-component signal. An ECG signal is a multi-component signal, which consists of the pulse frequency region (varying between about 1.2 and 1.4 Hz for the example shown in Figure 4.4, first panel) and its harmonics, which are located in regions given by integer multiples of the pulse region. The power of the EDA (Figure 4.4, second panel) and respiration (Figure 4.4, third panel) signals is primarily concentrated in one region, with only a small amount of power in other regions. Unlike the ECG and respiration signals, the EDA, in general, does not have a cyclic nature, but contains trends and abrupt changes. Most of its signal power is concentrated in the lowest frequencies (see Figure 4.4, second panel) and forms the delineated region of the EDA in the time-frequency plane. However, analysis of the EDA data showed that the EDA signal also contains power in other regions during some time intervals. For example, frequency domain analysis shows that both the EDA data of Figure 4.5 (first panel) and the synchronously measured respiration signal Figure 4.5 (second panel) contain power in the region around 0.4 Hz. Coherence analysis can reveal whether this is a coincidence. An interesting question is if or when this coherence occurs and whether the method of coupling depends on an emotional state. The existence of a small amount of power in higher frequency regions of the EDA signal, is evident in Figure 4.6 where the element-wise logarithm of the time-frequency distribution of the EDA signal depicted in Figure 4.4, second panel, is shown. For this example, the pulse frequency region and its harmonics become clearly visible in the EDA signal. For a thorough analysis of EDA signals, the interested reader is referred to (Lim et al., 1997) .

Figure 4.7 plots the pairwise coherences (as given by Equation (4.2)) of the three signals. For this example, the coherence between the ECG and EDA signals (first panel) is 'high' in the pulse frequency region and its harmonics, while it is 'low' in the frequency region that contains most of the EDA signal power. The coherence between the ECG and respiration signals (second panel), on the other hand, takes maximal values at the respiration frequency region, is 'moderate' at the pulse frequency region and its harmonics, and 'low' elsewhere. The coherence between the EDA and the respiration signals (third panel) is low for the respiration and EDA frequency regions and 'moderate' elsewhere. Instead of a narrative description of the pairwise relationships, we briefly describe a scalar measure which quantifies this information. For this, the concept of coherences within delineated regions in the time-frequency plane, as detailed in Muma et al. (2010), is followed. Figure 4.8 shows the pulse and respiration frequency regions. Muma et al. (2010) proposed an algorithm to detect these regions within the time-frequency plane. The idea of the algorithm, in brief, is: First, the frequency of the auto-spectrum that contains maximal energy $f_{\max,0}$ is determined for a data segment, in our case, of length 10 seconds. The frequency of maximal energy $f_{\max}(t)$ of the non-stationary signal varies around $f_{\max,0}$ and can be found by searching near the maximal value of the periodogram at each time instant. The delineated regions include all neighboring frequencies for which the power drops less than 3 decibel compared to $f_{\max}(t)$. By applying a mask onto $C_{XY}(t, f)$, which is equal to one within the detected regions

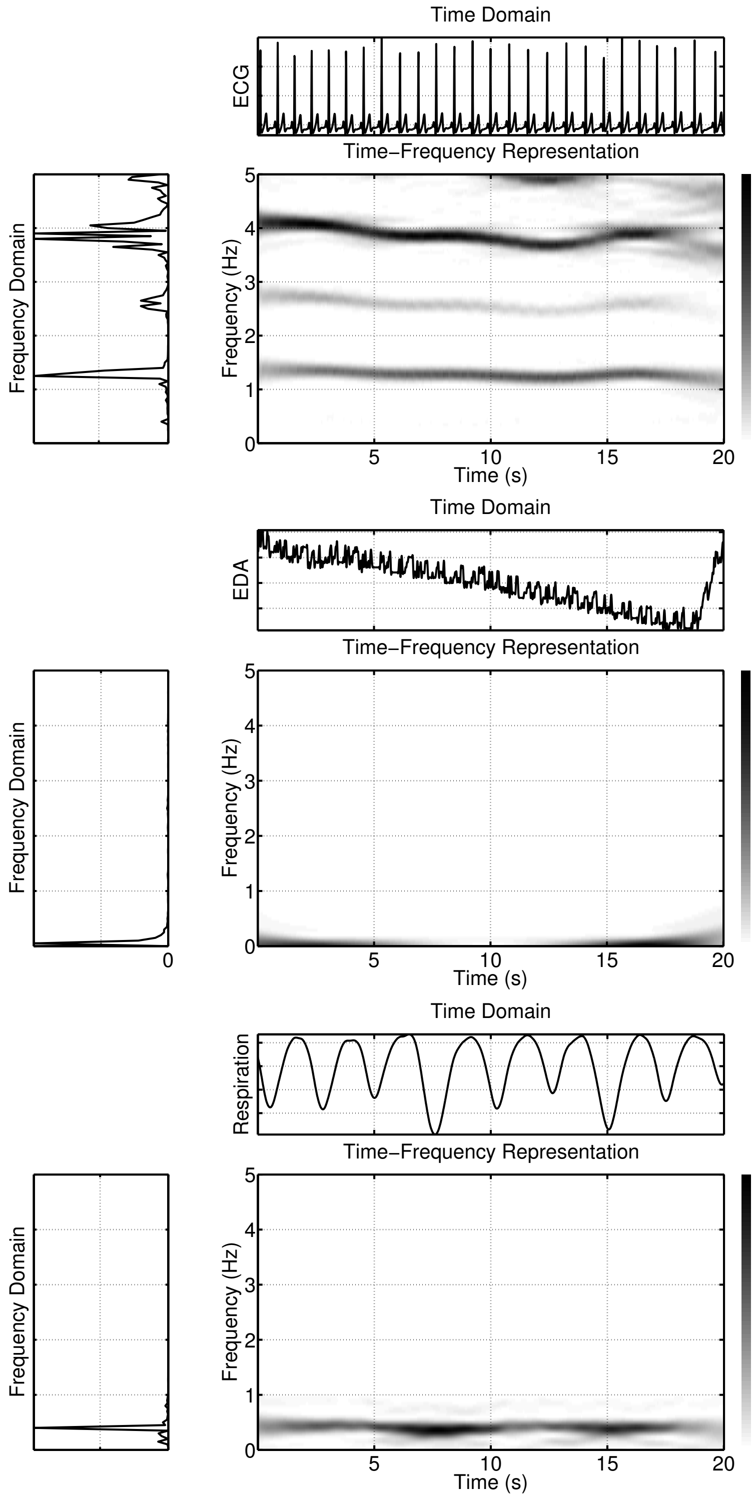


Figure 4.4. Example of ECG, EDA and respiration signals in time, frequency and time-frequency domains.

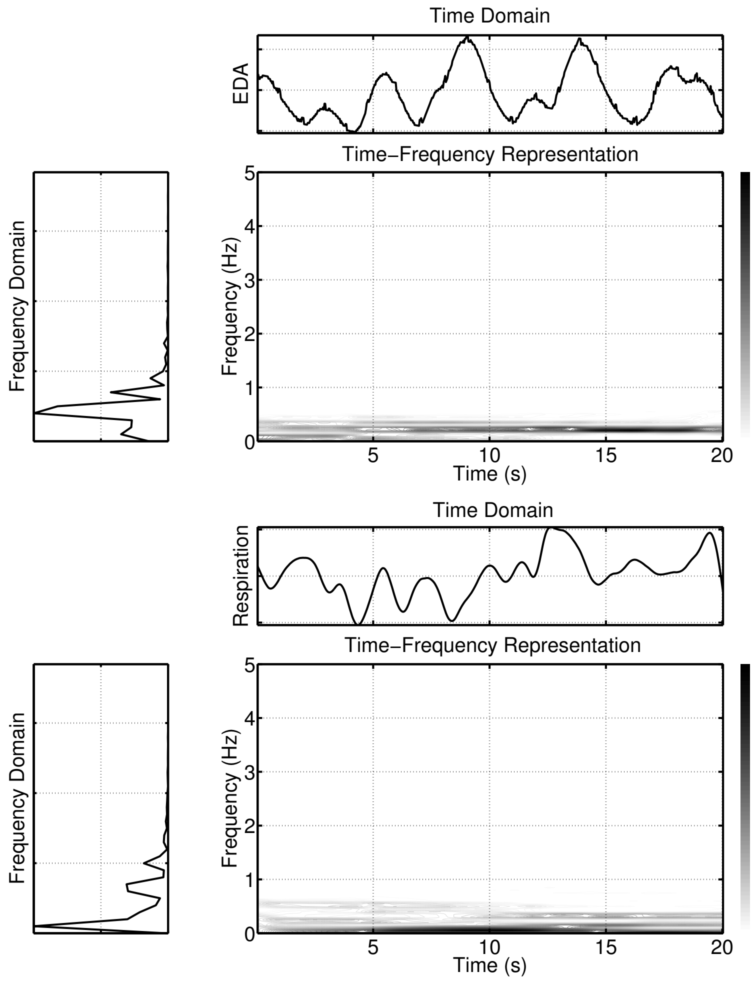


Figure 4.5. Example of synchronously measured EDA and respiration signals that both contain a cyclic component at the frequency region of approx. 0.4 Hz.

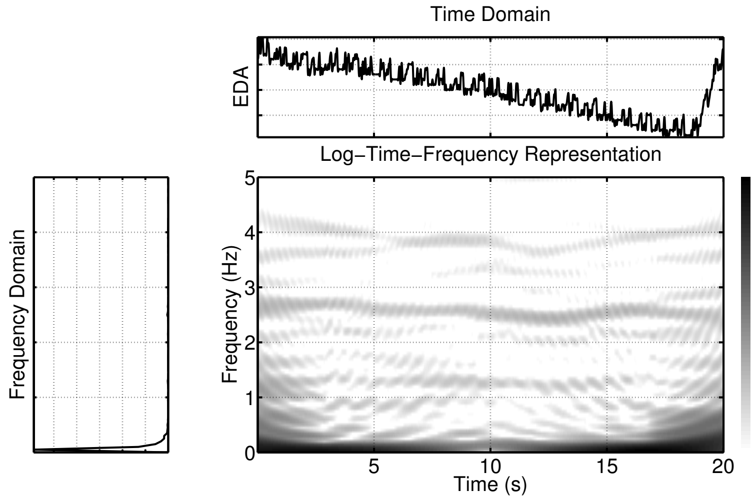


Figure 4.6. Element-wise logarithm of the time-frequency distribution of the EDA signal depicted in Figure 4.4, second panel. For this example, the pulse frequency region and its harmonics become clearly visible in the EDA signal.

and zero elsewhere (see Figure 4.8), analysis is restricted to the regions of interest. Note that, unlike classical spectrum based heart rate variability (HRV) analysis, (e.g., Niskanen, Tarvainen, Ranta-Aho, & Karjalainen, 2004; Acharya, Joseph, Kannathal, Lim, & Suri, 2006) and references therein, fixed frequency bands (VLF, LF, HF), which are set a-priori and remain constant over time, are not assumed. Our algorithm automatically adapts the frequency region of interest over time to the data at hand. Based on this, scalar time-varying coherence measures between two signals, such as, for example, the average coherence at a frequency region of interest for each time instant, can be determined. Figure 4.9 plots the pairwise (bivariate) coherence measure for the three signals in the respective frequency regions. It is clearly visible from these examples that a coupling between the signals exists and that the amount of synchronization varies over time. In the next section, a system-wise coherence measure based on the pairwise coherences, is defined.

A latent state space modeling approach to a system-wise synchrony measure

In this subsection, the well-known state space modeling approach (e.g., Durbin & Koopman, 2001) is used to specify an overall synchrony measure. The basic idea is to use the multivariate time series of pair-wise coherence measures within delin-

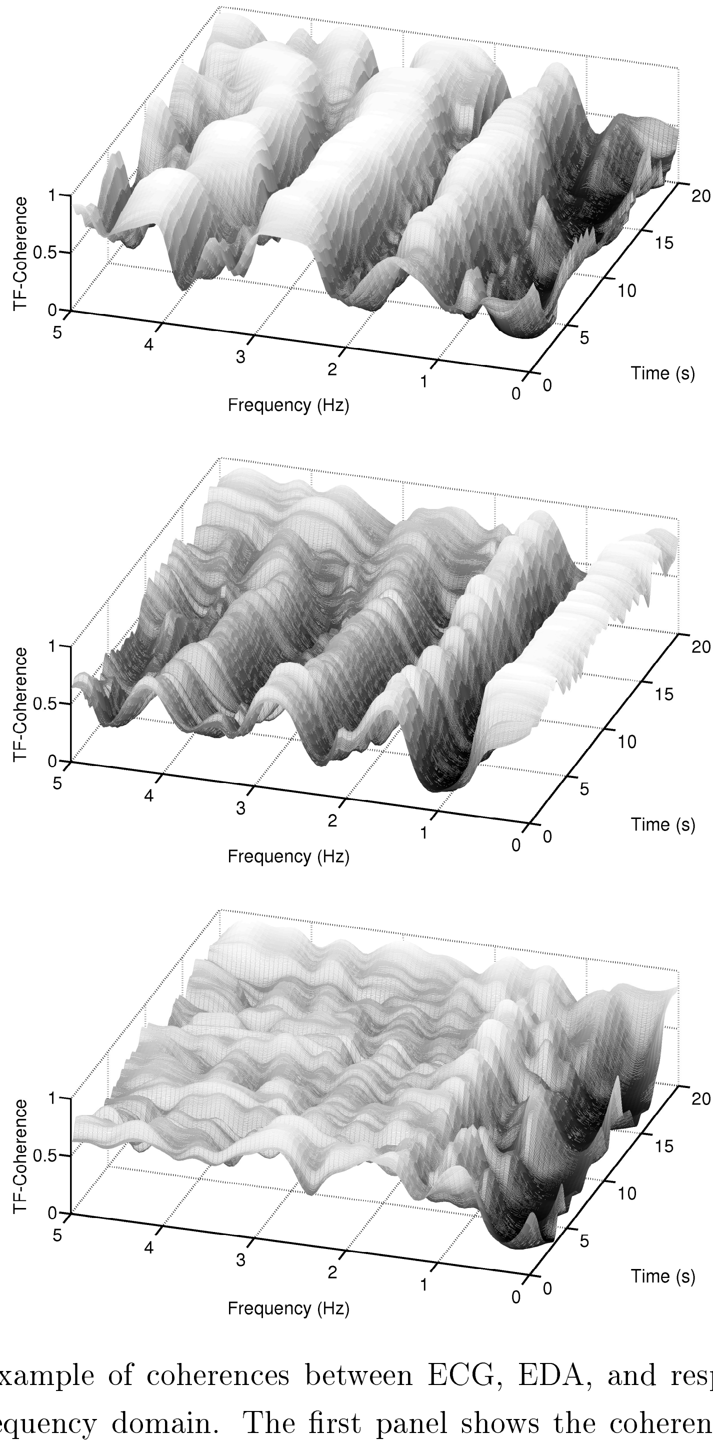


Figure 4.7. Example of coherences between ECG, EDA, and respiration signals in the time-frequency domain. The first panel shows the coherence between the ECG and EDA signals, the second panel shows the coherence between the ECG and respiration signals, and the third panel shows the coherence between the EDA and respiration signals.

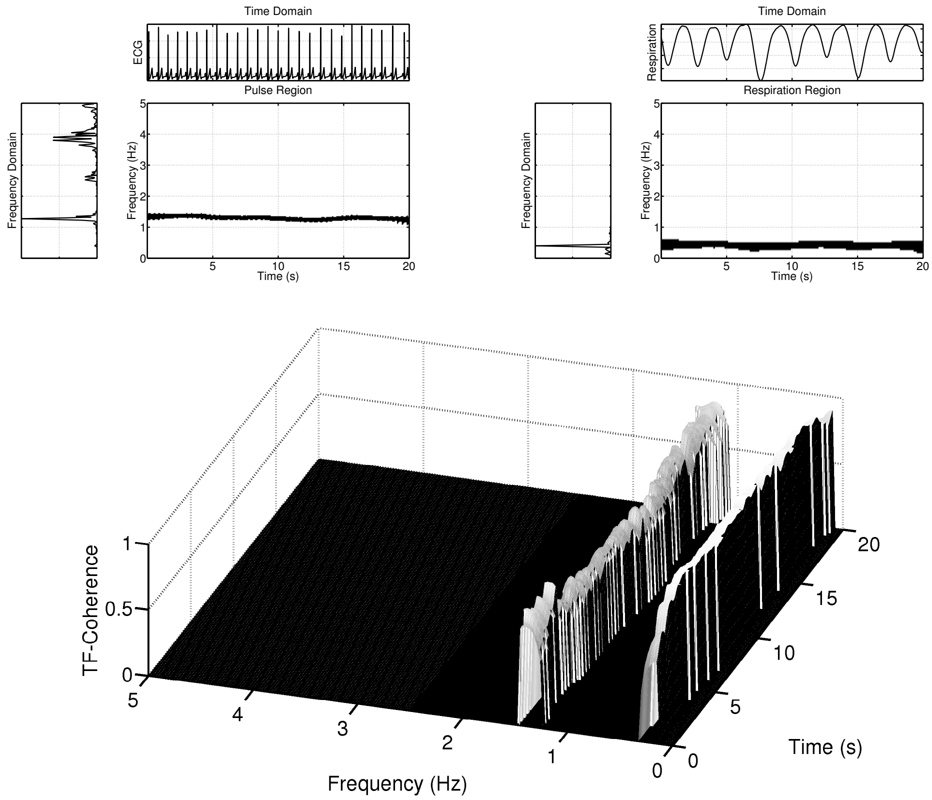


Figure 4.8. Example of coherences within delineated regions, i.e. the pulse and respiration regions, between ECG and respiration signals.

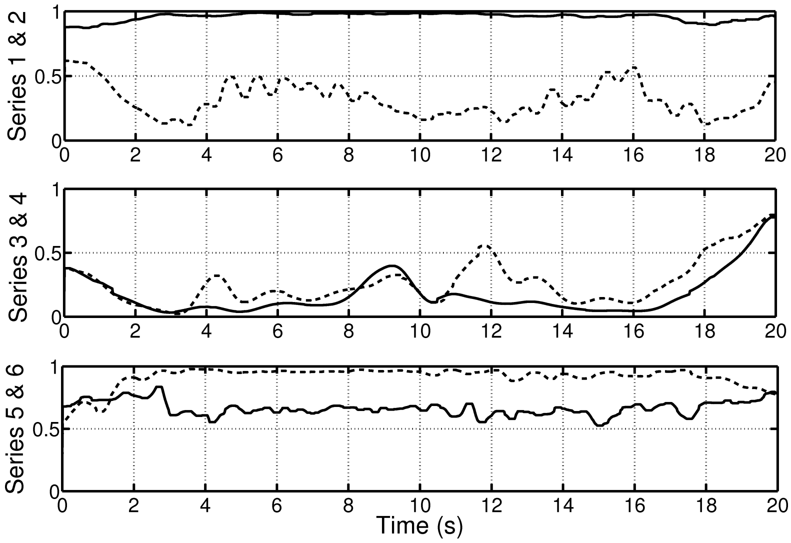


Figure 4.9. Example of the pairwise coherence measure based on the concept of delineated regions in the time-frequency plane as shown in Figure 4.8. Series 1 & 2 plots the coherence of the ECG and EDA signals in the pulse region (solid) and EDA region (dashed), respectively. Series 3 & 4 depict the coherence of the respiration and EDA signals in the respiration region (solid) and EDA region (dashed). Series 5 & 6 depict the coherence of the respiration and ECG signals in the pulse region (solid) and respiration region (dashed).

eated frequency regions (see previous subsection and Figure 4.9) and to specify a measurement model that operationalizes a latent variable which represents an underlying state of an overall system-wise synchrony. In addition to the measurement model, a structural model describes the regression of a state relative to its previous states.

For a given individual i, the specification of the measurement and structural model is given in the so-called non-innovation form (e.g., P. D. Gilbert, 1993):

$$z_{it} = H_t \xi_{it} + R_t e_{it} \tag{4.3}$$

$$\xi_{it} = F_t \xi_{i(t-1)} + G_t u_{it} + Q_t \eta_{it} \tag{4.4}$$

In the measurement model (Equation (4.3)), we assume that a given p-dimensional observed variable z_{it}, at time point t, can be regressed on a q-dimensional (latent) state vector ξ_{it}, where e_{it} is a p-dimensional residual vector (white noise). The variable z_{it} hereby includes the pair-wise coherence measures in selected frequency regions (delineated regions), whereas ξ_{it} represents the overall synchrony for an individual i at time point t, which is the variable of interest. The estimation of the pair-wise coherence measures is described in the previous section. H_t and R_t are time-varying coefficient matrices, respectively of dimensions $(p \times q)$ and $(p \times p)$ that reflect the time-varying relationship between the latent variable vector ξ_{it} and the observed variable vector z_{it}, and the time-varying relationship between the residual vector e_{it} and the observed variable vector z_{it}, respectively. In the structural model (Equation (4.4)), it is assumed that the state vector ξ_{it} can be regressed on a previous state vector $\xi_{i(t-1)}$ and on a m-dimensional covariate vector u_{it} (e.g. an intervention), where η_{it} is a q-dimensional latent residual (white noise) vector. Again, $F_t(q \times q), G_t(q \times m)$, and $Q_t(q \times q)$ are time-varying coefficient matrices. The coefficient matrix H_t indicates the time-dependent strength of the relationship of the state variable ξ_{it} and its indicators/measures.

Predictions of latent states ($\xi_{i(t+1)}$) can be conducted, for example, by applying the Kalman filter (e.g. Grewal & Andrews, 2001; Shumway & Stoffer, 2006). By using a smoothing algorithm (e.g. Durbin & Koopman, 2001), scores for the latent variable ξ_{it} can be derived for the time series. The model parameters can be, for example, determined via a maximum likelihood estimator and imposing additional distributional assumptions (for example, normality of the latent variable ξ_{it}). It is not possible to freely estimate all the parameters in Equations (4.3) and (4.4).

In order to have an identified model, the typical additional restrictions of latent variable modeling (e.g. scaling variables) are needed (Bollen, 1989). For a given individual i the time-dependent pattern of the parameters reflects the changing structural relationship of the individual reactions in the observed variables over the course of time; specifically, here, the psychophysiological responses (e.g. Marwitz & Stemmler, 1998). The time series of the resulting latent states ξ_{it} represents the course of the overall system-wise synchrony.

Figure 4.10 gives an example of a time series of the system-wise synchrony measure (latent state variable ξ_{it}) during an emotional episode. The first ten seconds (100 points) show the synchrony while a participant was watching a affective neutral picture (barstool: #7025) of the International Affective Picture System (IAPS; Lang et al., 2008). During the last ten seconds (100 points) the participant was confronted with a disgust eliciting picture (half ripped-off finger: #3150). As a descriptive result, it can be seen the overall synchrony measure was, on average, higher during the disgust eliciting picture than during the neutral picture.

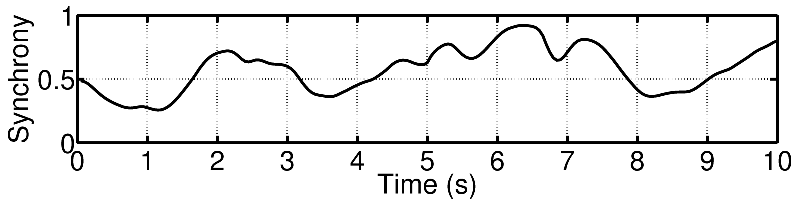


Figure 4.10. Example of an overall synchrony measure while a participant was watching a neutral picture and a disgust eliciting picture for 10 seconds each.

4.3. Application of Two Approaches for the Quantification of Synchrony Using Data from an Emotion Regulation Study

In this section, we illustrate the application of the proposed measure of synchrony, and the application of a measure proposed by Hsieh et al. (2011), to data from a larger emotion regulation study where psychophysiological signals were collected while participants were watching funny film clips. First, a short description of the larger study, is given. Second, assuming that the induction of emotions leads to person-specific changes in synchrony of the psychophysiological signals, the time series of the overall synchrony measure is analyzed by applying the proposed approach. Third, the results are compared with those obtained by using the approach for the quantification of the (aggregated) synchrony of Hsieh et al. (2011).

Description of the emotion regulation study

Participants The sample of the emotion regulation study consisted of 58 German male undergraduate students of engineering. The mean age of the overall sample was 23.11 years (SD = 3.72). The procedure was fully explained to participants before written informed consent was obtained. Participants took part in a lottery where they could win a portable media player.

Procedure Participants watched a random sequence of two funny clips (each ten-minutes long), which were cuts of the German version of the slapstick comedy film "You Don't Mess with the Zohan", starring Adam Sandler. In a small pilot-study, ratings of the clips showed that one film clip was slightly funnier than the other. Blood pressure was measured five and ten minutes after each clip. Physiological measures were obtained continuously, and self-reported ratings of experienced amusement were collected for each film clip.

Measures The data consisted of physiological and subjective responses during the film clips. The subjective responses are not considered here. Physiological measures were continuously sampled with a frequency of 256 Hz. The following signals were obtained using a BIOPAC MP150 System (Biopac Systems Inc, 2011): ECG, EDA, and respiration signals. In addition to these continuous signals, discrete measures of systolic and diastolic blood pressure were taken every five minutes. In this article, analysis was restricted to the ECG, the EDA, and the respiration signal. Motion artifacts in the ECG signals were removed using a method proposed by Strasser, Muma, and Zoubir (2012).

Application of the proposed approach for the quantification of synchrony

In order to obtain an overall quantity of the synchrony, the following two step procedure was applied. First, the time-frequency based pairwise coherence measures were obtained. Using three signals (here: ECG, EDA, and respiration) results in six pairwise measures (two measures for each pair of physiological signals; see for example Figure 4.9). Figure 4.11 gives an example of the six resulting time series for one participant over a period of 120 seconds. During the first 60 seconds, the participant was watching the film clip, which was rated as moderately funny. During the last 60 seconds, the participant was shown the funnier film clip. As can be seen from Figure 4.11, some of the pairwise measures are very similar (for example Series 5 and 6).

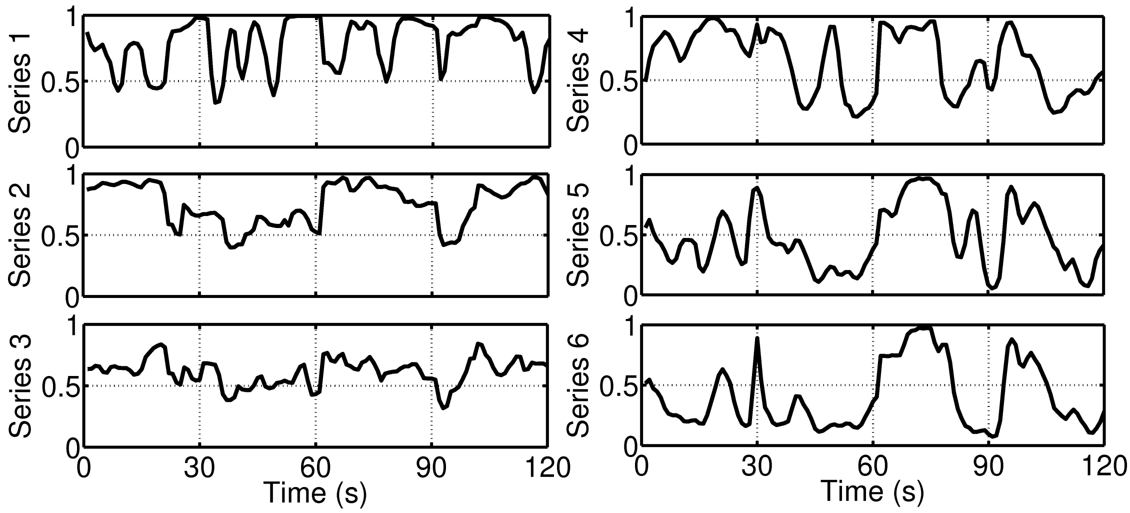


Figure 4.11. Example of six time-frequency based pairwise coherence measures (series) for one individual watching funny film clips. Series 1: ECG-EDA coherence (EDA region), Series 2: ECG-EDA coherence (ECG region), Series 3: RE-ECG coherence (ECG region), Series 4: RE-ECG coherence (RE region), Series 5: RE-EDA coherence (EDA region), Series 6: RE-EDA coherence (RE region). During the first 60 seconds, the participant was watching a film clip, which was rated as moderately funny. During the last 60 seconds, the participant was shown a funnier film clip.

Second, the obtained pairwise measures $z_{it} = (z_{1it}, z_{2it}, ..., z_{6it})^{\mathrm{T}}$ were indicators of a single overall latent variable ξ_{it} time series of the system-wise synchrony using the state-space modeling approach as described above (cp. Equation (4.3)). For a given person i, the measurement variable vector is given as:

$$z_{it} = (z_{1it}, z_{2it}, ..., z_{6it})^{\mathrm{T}} = (a_{1i}, a_{2i}, ..., a_{6i})^{\mathrm{T}} + (h_{1i}, h_{2i}, ..., h_{6i})^{\mathrm{T}}\xi_{it} + (e_{1it}, e_{2it}, ..., e_{6it})^{\mathrm{T}} \tag{4.5}$$

where $a_{1i}, ..., a_{6i}$ represent additional intercepts, $h_{1i}, h_{2i}, ..., h_{6i}$ are loadings, and

$e_{1it}, ..., e_{6it} \sim N(0, \sigma_e)$. $N(0, \sigma_e)$ is the zero Gaussian distribution with a standard deviation equal to σ_e. The latent variable ξ_{it} was assumed to be uni-dimensional (see Equation(4.4)), reflecting the overall synchrony. For a given person i, the structural model is given as:

$$\xi_{it} = g_{1i} + f_{1i}\xi_{i(t-1)} + \eta_{it} \tag{4.6}$$

where g_{1i} is an intercept, f_{1i} is a regression coefficient, and $\eta_{it} \sim N(0, \sigma_\eta)$. Figure 4.12 provides examples of overall synchrony measures for two participants. Again, during the first 60 seconds, participants were watching a film clip, which was rated as moderately funny. During the last 60 seconds the participants were shown a funnier film clip.

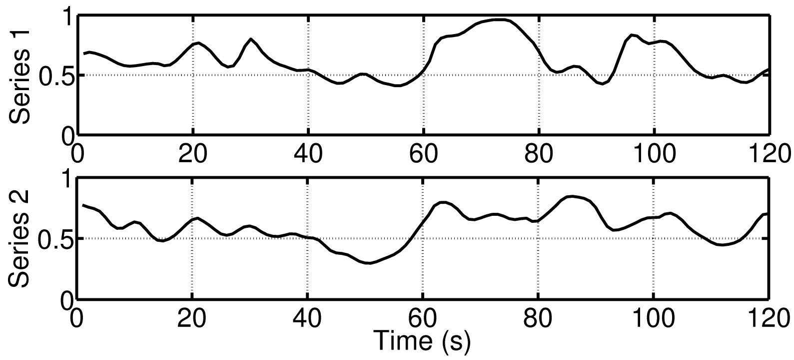


Figure 4.12. Example for two overall synchrony measures for two participants. During the first 60 seconds, the participants were watching a film clip, which was rated as moderately funny. During the last 60 seconds, the participants were shown a funnier film clip.

Table 4.1 presents parameter estimates, standard error estimates, and confidence intervals for two participants (see Equations (4.5) and (4.6)). The results between the two participants vary substantially. As can be seen from Table 4.1, for participant #1 the loadings h_1 and h_3 are not significant and thus the measures z_1 and z_3 are not reliable indicators of synchrony. The other indicators (z_2, z_4, z_5, z_6) indicate the degree of overall synchrony. In contrast, for participant #2 a different set of indicators (z_1, z_4, z_5, z_6) imply an overall level of synchrony. Furthermore, the sign of the loading of indicator z_4 is different, which means that for participant #1 an increase in the overall synchrony leads to an increase in the relationship of the two variables associated with respiration and EDA ($h_4 = 0.97$), while for participant #2 an increase in the overall synchrony leads to an decrease in the relationship of the two variables associated with respiration and EDA ($h_4 = -0.39$). These simple examples show that the patterns of synchrony of the psychophysiological signals are very person-specific.

Table 4.1.

Parameter Estimates (Est.), Standard Errors (SE) and 95%-Confidence Intervals (CI) of the Parameter Estimates Obtained From the Estimated Latent State Space Model Approach for the Quantification of an Overall Measure of Synchrony. During the First 60 Seconds, the Participants Were Watching a Film Clip, Which was Rated as Moderately Funny. During the Last 60 Seconds, the Participants Were Shown a Funnier Film Clip.

	Participant #1				Participant #2			
	Est.	SE	low.CI	up.CI	Est.	SE	low.CI	up.CI
h_1	0.0166	0.1157	-0.2283	0.2320	1.1227	0.1289	0.8430	1.3720
h_2	0.3134	0.1146	0.0795	0.5366	-0.2842	0.1435	-0.5420	0.0113
h_3	0.2358	0.1144	-0.0162	0.4614	-0.0065	0.1526	-0.2967	0.2801
h_4	0.9739	0.0988	0.7768	1.1625	-0.3869	0.1470	-0.6510	-0.0572
h_5	1.4766	0.0924	1.2625	1.6341	1.3736	0.1205	1.1132	1.5821
h_6	1.5386	0.0897	1.3394	1.6910	1.4772	0.1185	1.2175	1.6918
a_1	0.7722	0.0731	0.6328	0.9305	-0.0278	0.0846	-0.1662	0.1657
a_2	0.5616	0.0730	0.4189	0.7109	0.7988	0.0854	0.6161	0.9467
a_3	0.4650	0.0714	0.3289	0.6155	0.5641	0.0911	0.3933	0.7405
a_4	0.0404	0.0654	-0.0762	0.1814	0.9493	0.0888	0.7423	1.1042
a_5	-0.4537	0.0635	-0.5592	-0.2991	-0.3172	0.0801	-0.4389	-0.1175
a_6	-0.5598	0.0603	-0.6554	-0.4261	-0.4451	0.0818	-0.5591	-0.2404
σ_e	0.0225	0.0012	0.0198	0.0248	0.0285	0.0016	0.0251	0.0313
f_1	0.9243	0.0566	0.7389	0.9580	0.9145	0.0578	0.7267	0.9585
g_1	0.0462	0.0343	0.0254	0.1575	0.0499	0.0325	0.0221	0.1488
σ_η	0.0035	0.0010	0.0018	0.0056	0.0028	0.0009	0.0012	0.0050
ξ_0	0.6832	0.0968	0.5025	0.8725	0.7912	0.0945	0.6101	0.9928

Note. Confidence intervals were obtained using a (parametric) bootstrap procedure provided by the MARSS package in R-project.

Application of a complementary approach for the quantification of the synchrony proposed by Hsieh et al. (2011)

To provide a comparison to existing procedures, the recently proposed method by Hsieh et al. (2011) is applied to the dataset. The method estimates an aggregated system-wise synchrony by creating a stochastic network, where high connectivity corresponds to high synchrony of the $J = 3$ signals (here: ECG, EDA, and respiration signals).

First, receiver operating characteristics (ROC) were obtained. For a random variable X, the ROC displays the probability of correct detection, i.e. the probability of correctly accepting the null hypothesis, as a function of x, versus the probability of false alarm, i.e. the probability of falsely accepting the null hypothesis, as a function of x. In our experiment the null hypothesis (H_0) is associated to the time interval with the moderate funny film clip, i.e. the baseline, while the alternative (H_1) is associated with the funnier (emotionally more activating) film clip. The ROC area A is defined as the area between the ROC curve and a diagonal line given by an identity function:

$$A = \int_0^1 (P(H_0|H_0, x) - c(x)) \frac{dP(H_0|H_1, x)}{dx} \, dc(x).$$

Here, $c(x)$ is a linear function, that maps the interval of x onto the interval $[0, 1]$, which is equivalent with the diagonal line in this context. $P(H_0|H_0, x)$ and $P(H_0|H_1, x)$ are the probabilities of correct detection and false alarm given a threshold at x, respectively (Hsieh et al., 2011). For the signals acquired while the participants were watching both film clips, in accordance with Hsieh et al. (2011), are divided into $\frac{K}{2} = 12$ non overlapping segments. The ROC areas were then calculated by defining $P(H_0|H_0)$ as the probability of deciding for a particular film clip given that the participants were actually watching this film clip. In this case, we used the cumulative distribution function (CDF) of the single segments, and $P(H_0|H_1)$ as the probability of deciding for a particular film clip given that the other film clip was watched. $P(H_0|H_0)$ and $P(H_0|H_1)$ are used to form the CDFs of the baseline and of the activation phase, respectively.

In the second step, the Spearman rank coefficients of the ROC areas of 24 non-overlapping blocks were calculated. These were formed by 12 blocks of 50 seconds each from the emotionally neutral phase and 12 blocks of 50 seconds each from the emotionally eliciting phases. It should be noted that the acquired signal blocks exhibited a non-stationary character. By calculating the Spearman rank coefficients of the ROC area sequences of the non-stationary signal blocks for each individual and signal-pair, a single scalar value is obtained as an averaged measure of coherence. This scalar does not take into account any stochastic changes of the signals, and, hence, it disregards significant information. Applying this procedure to every subject and signal pair, yields a $J \times J$ rank coefficient matrix for each subject and phase. Applying

$$\hat{p}(h|j,j') = \frac{1}{M} \max \left\{ \sum_{m=1}^{M} 1_{\{\hat{\rho}^{(m)}(j,j')>h\}}, \sum_{m=1}^{M} 1_{\{\hat{\rho}^{(m)}(j,j')<-h\}} \right\} \tag{4.7}$$

from Hsieh et al. (2011) with the critical value $h = 0.344$ taken from the Spearman rank distribution table (see Spearman; for $n = 24$ and $\alpha(1) = 0.05$), yields a $J \times J$ matrix, called the partition matrix:

$$\hat{p}(h) = \begin{bmatrix} - & 0.3820 & 0.4720 \\ 0.3820 & - & 0.3146 \\ 0.4270 & 0.3146 & - \end{bmatrix},$$

Here the 1st, 2nd, and 3rd rows and columns belong to the ECG, EDA, and respiratory signals, respectively.

In the third step, a stochastic network was defined, where the nodes correspond to the ECG, EDA, and respiratory signals, and the edges were defined by values from the partition matrix $\hat{p}(h)$. To introduce randomness, a matrix u was formed whose entries were generated by the uniform distribution on the interval $[0, 1]$. In general, an edge between two nodes j, j' was established, if $b_{\text{con}}(j, j') = p(j, j') > u(j, j')$, i.e., if the Boolean matrix b_{con} had a '1' entry at (j, j'). If $p(j, j') \leq u(j, j')$, there was no connection between j and j' nd the correspondent entry in the b_{con} matrix was 0. Each node of the stochastic network is either activated or inactive. To test for system-wise synchrony, in the beginning, a node must be selected for activation. After activation, the node sends a signal, via its edges, to other directly connected nodes, which then activate their neighboring nodes. After sending out the signal, the node that was activated in the first place becomes inactive. This activation, and inactivation, of nodes was iterated until (i) all nodes became activated, in which case system-wise synchrony was achieved, or (ii) a system state was obtained for which system-wise synchrony was impossible to achieve. For J nodes, this was done J times, each time activating a different node at the beginning. In our case of $J = 3$ nodes, system-wise synchrony could only be achieved by a fully connected network. Based on a Monte Carlo simulation with 10,000 repetitions, the stochastic network achieved system-wise synchrony 511 times. Thus, the resulting stochastic network had a probability of about 5.1% to achieve system-wise synchrony.

Due to the previous averaging over the subjects, the calculated probability gives us an expected value for achieving system-wise coherence for any of the given subjects. It is important to realise that the result of the simulation is not dependent on an individual subject and that the results depend on the assumption of stationarity.

4.4. Discussion

A new approach for the quantification of synchrony of multivariate non-stationary psychophysiological signals has been proposed. After calculating bivariate time-frequency based coherence measures, a state space modeling procedure was applied to obtain an overall measure of synchrony. The approach gives information on the intraindividual level about the course of synchrony of psychophysiological reactions.

Methodological and substantive considerations

The use of multivariate time series of physiological reactions for the quantification of an overall synchrony has several methodological and substantive implications. Firstly, the approach provides time-sensitive information about the intraindividual level of synchrony. Thus, it is complementary to alternative approaches, which give information on cross-subjects based (aggregated) synchrony. See, for example the approach proposed by Hsieh et al. (2011)[5].

[5] On p. 143 of their publication, Hsieh et al. write: "There is a growing body of psychological literature questioning whether aggregates are useful descriptions of the individual (e.g., Hamaker, Dolan, & Molenaar, 2005; Molenaar, 2004; Nesselroade & Molenaar, 1999. Our ROC curve analyses were conducted at the individual level and this information was then compiled in all subsequent analytic steps. Although perhaps not an optimal approach, our analyses were built upon individual information."

Secondly, the proposed approach also provides information about the interindividual difference of the structure of the synchrony measure. Specifically, the person-specific reaction patterns of how the physiological signals are related with the (latent variable) synchrony have been quantified. The factor loadings within the state space approach provide person-specific information on how the simultaneous activation of two given signals, e.g. ECG and EDA signals, is determined by an overall synchrony and whether the same signals are reliable indicators of synchrony across individuals. As we have seen from the above discussed empirical example, this is not necessarily the case.

Thirdly, by applying the time-frequency based procedure from Muma et al. (2010), the proposed approach directly addresses problems of non-stationarity. In general, non-stationarity is a challenge in the examination of time series (Scharf, 1991; Vaseghi, 2008). When a relationship between longitudinal data has to be quantified, non-stationarity leads to a underestimation of the time-dependent relationship (see examples from above).

Fourthly, a practical implication of the proposed approach is that researchers are given a simple two step tool, which is able to answer foundational research questions from emotion theory. The (real time) temporal resolution of synchrony of (non-stationary) measures allows for a person-orientated identification of response systems that have changing distributional characteristics over time (for example respiration rate). By addressing substantive questions on the intraindividual level the construct validity of concepts based on psychophysiological data can be strengthened (Borsboom, Mellenbergh, & van Heerden, 2003).

Lastly, there are implications for emotion theory. The proposed approach facilitates assessing whether emotion eliciting stimuli lead to a stronger synchrony of psychophysiological reactions during positive and negative emotional states. In principle, within the time series representation of the state space model, treatment effects can be specified. With our empirical data collected from a study on humor we only found weak (descriptive) effects between two funny film clips. However, recent results from other studies on disgust indicate that differences between phases can be found. Based on our research, it is possible to conclude that testing for treatment effects in a time series is a precise and powerful tool which is in some cases superior to examining treatment effects with aggregated data.

Limitations and future directions

The proposed approach provides novel possibilities for analyzing psychophysiological signals, nevertheless it also has some limitations. Both are briefly discussed: Although being able to incorporate non-stationary signals, the proposed approach is not capable of analyzing signals of constant amplitude. For example, when subjective data are measured using a potentiometer, participants can continuously adjust their subjective experience during emotional stimuli. Unfortunately, participants do not change their ratings continuously. Between changes in the position of the potentiometer, the signals are constant. For such time intervals, the proposed approach can not be used for the quantification of synchrony of physiological and subjective data. Nevertheless, participants have an enduring/varying emotional experience in this period. This general problem, which also occurs for (retrospective) paper-and-pencil ratings, remains unsolved. A potential solution could be an extension of the proposed approach, which operates on a feature space of the signals. Specifically, instead of looking directly at the bivariate coherence of measured signals, a feature estimation step could be added that extracts psychophysiological features from the signals (on a higher data level). The bivariate coherence from the first step may be based on the synchrony of these features. Enhanced robustness and a higher flexibility would possibly be the result. Also utilizing the feature space may save computational resources: Features change more slowly, which allows for reducing the sampling frequency in the coherence computation.

An additional limitation of the proposed approach is the assumed linearity of the bivariate coherences. In future work, the linear coupling between the signals could be relaxed, which would reflect a more realistic specification of underlying processes.

A possible future research direction consists in a detailed examination of coherence, and overall synchrony patterns, for specific individuals during specific emotion eliciting situations. On the one hand, individuals and their emotions could be recognized in applied settings (e.g., man-machine interaction). On the other hand, basic research questions of emotion theory could be addressed on a more detailed level. For this, further evaluation of more psychophysiological signal types in different emotion eliciting situations will be necessary. For example, the proposed approach works well in situations, in which emotions are elicited during a relatively long period of time (here several seconds). However, it is unclear how the proposed approach performs in the case of, for example, emotional priming, in which emotions are elicited for a very short period of time. The applicability of the proposed approach strongly depends on the responsiveness of physiological signals that were chosen for the analysis. Respiration, ECG, and EDA, for instance, are very slow or inadequate for the representation of affective responses in the context of priming. Therefore adequate signals such as EEG waves should be chosen for the analysis.

From a practical perspective, it would be interesting to extend the use of the proposed approach to applications in the context of the strongly developing research field of affective computing (e.g., Picard, 1995, 1997; Scherer, Banziger, & Roesch, 2010). Computers, which recognize, interpret, and process emotions could be used not only in health services, but also in education science, human-computer-interaction and other applications. Regarding the performance of the affect detection, the comparison of different algorithms (in terms of classifiers), would be an important field of research (Hudlicka, 2003; Kolodyazhniy et al., 2011), which also depends on the integration of different measured response signals, for example speech, physiological, and behavioral/mimic measures. Clearly, response signals differ in the sense that some are more informative than others, depending on which emotions need to be separated (e.g., Larsen, Norris, & Cacioppo, 2003).

Furthermore, as an important application, biofeedback might be considered (Thompson & Thompson L., 2003). Since biofeedback addresses numerous signal types, such as behavior, muscle tone, brainwaves, breath, skin conductance, heart rate, temperature and pain perception, and many more, an adaptation of the proposed approach needs to be conducted in future research in order to be applicable in different settings (e.g., Dawson, Schell, & Filion, 2007; Tassinary, Cacioppo, & Vanman, 2007; LaVaque, 2003).

Finally, although beyond the scope of this article, an interesting field of research has emerged that examines the synchrony of psychophysiological signals and empathy (Hulsman, Smets, Karemaker, & de Haes, 2011; Marci & Orr, 2006; Oliveira-Silva & Gonçalves, 2011; Reed, Randall, Post, & Butler, 2013). In this research, couples of subjects have been examined with respect to their synchrony of psychophysiological responses in the context of empathy (Levenson, 1992). Although, the neurobiological mechanisms underlying the synchronous empathic responses are mostly unclear, an adaption of the proposed approach might be helpful for further research in the field. The extension is not straightforward, because the interindividual level is added to the current setting. Therefore, it might be interesting from a psychometric perspective to examine the synchrony of responses that stem from related but not identical subjects.

5. Manuscript B: Synchrony of Psychophysiological Parameters in Disgust

Abstract

The synchrony of physiological reactions constitutes an inherent part of many emotional theories. Recently, new quantification methods have been introduced but the temporal characteristics of a physiological synchrony response are largely unknown. Our aim was to explore short-term intraindividual changes in the physiological synchrony of selected heart, respiratory, and skin conductance parameters in disgust. Thus, we focused on the beginning, time course, end, and level of physiological synchrony in reactions to disgusting versus neutral pictures. The temporal description of a peripheral physiological synchrony response is novel in itself. We applied Kelava, Muma, Deja, Dagdagan, and Zoubir's (2015) new quantification procedure which accounts for the complex data structure of multivariate and non-stationary input variables for which the probability distributions of variables change over time. We also considered the synchronously acquired continuous self-reported intensity of disgust measured with a rating dial. The sample consisted of 42 participants, who watched neutral and disgusting pictures. On average, physiological synchrony began to increase in a similar fashion for neutral and disgusting pictures while the participants viewed the black screen that preceded a new picture. However, physiological synchrony was significantly higher when a disgusting picture appeared compared with a neutral picture. In addition, the average physiological synchronization was faster and shorter than subjective ratings of disgust. In line with current empirical research results, large interindividual differences in the amplitude, slope, and duration of physiological synchrony were observed. The results

are discussed in the light of emotion theories and embedded in current empirical research.

5.1. Introduction

In the past, many theoretical approaches have been proposed to explain the appearance of an emotion and the associated changes in the brain, cognition, body, behavior, and subjective experience (for an overview, see Moors, 2009). The approaches vary, amongst others, in their assumptions about the temporal order of the changes and the causal mechanisms of emotions (e.g., Barrett, 2006a; Lazarus, 1991; Panksepp, 1982; Schachter & Singer, 1962; Scherer, 2009; Tomkins, 1962). Furthermore, some approaches propose a distinct number of distinguishable emotional categories, such as fear and joy (e.g., Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962), whereas others propose that emotional experience can instead be explained through different dimensions, such as valence and arousal (e.g., Barrett, 2006a; Russell, 1980). These different theoretical assumptions result in different views about the functionality of synchrony in emotional response parameters, meaning the joint co-occurrence of physiological, behavioral, and subjective changes during an emotional experience[6] (e.g., Barrett, 2013; Clore & Ortony, 2013; Coan, 2010; Ekman, 1992; Levenson, 1994).

In this old but still ongoing debate (e.g., Lench et al., 2011; Lindquist, 2013), there are two main positions, which, for example, can be traced back to early statements made by Darwin (1872/1965) and James (1884). On the one hand, beginning with Darwin's point of view, which was picked up and developed further by representatives of the basic emotion approaches (e.g., Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962), a stimulus elicits an emotion, which, by activating certain brain circuits, causes emotion-specific changes in multiple response modalities (e.g., physiological parameters, behavior, and subjective experience). These changes should be synchronous and similar across individuals (Ekman, 1992). Some of the so-called appraisal approaches share similar ideas, with the difference that the

[6] Several terms have been introduced to describe these simultaneous changes in response variables: coherence (e.g., Dan-Glauser & Gross, 2013; Evers et al., 2014; Hsieh et al., 2011; Mauss et al., 2005; Rosenberg & Ekman, 1994; Sze et al., 2010), concordance (e.g., Bulteel et al., 2014; Butler et al., 2014; Friedman et al., 2014; Hollenstein & Lanteigne, 2014), synchrony (e.g., Grandjean et al., 2008; Kelava, Muma, Deja, Dagdagan, & Zoubir, 2015; Scherer, 2009), and the organization of response systems (e.g., Levenson, 1994). Throughout this paper, we use the term synchrony.

emotions are not evoked through a stimulus itself, but rather through the person's appraisal of the stimulus (e.g., Lazarus, 1991; Roseman et al., 1990). Recently, Coan (2010) incorporated these assumptions into a measurement model called *the latent variable model*. Here, the emotion as the latent exogenous variable causes changes in the indicator variables, such as neural activity (e.g., amygdala activity), physiological signals (e.g., change in heart rate), behavioral change (e.g., facial expression), and subjective experience (Coan, 2010). Thereby, the level of synchrony depends on the intensity level of the elicited emotion (Coan, 2010), which corresponds with the assumptions of the basic emotion approaches (e.g., Ekman, 1992; Levenson, 1994; Rosenberg & Ekman, 1994). From this point of view, the synchronous change serves as a biological function in terms of optimal preparation for action (Levenson, 1994; Rosenberg & Ekman, 1994).

On the other hand, beginning with James' point of view, which was picked up and developed further by Schachter and Singer (1962) and representatives of psychological construction approaches (e.g., Barrett, 2006a; Cunningham et al., 2013; Ortony et al., 1988; Russell, 2009), an emotion emerges through the process of evaluating incoming contextual, bodily, and cognitive information. There is no fixed number of distinct emotional categories; instead, emotions have some sort of affective core, such as valence (e.g., Clore & Ortony, 2013; Cunningham et al., 2013), arousal (e.g., Schachter & Singer, 1962), or valence and arousal (e.g., Lindquist et al., 2013; Russell, 2003). The situation, personal history, and experience all shape the affective core, which is then labeled, for instance, disgust, thereby allowing for inter- and intraindividual variability (Barrett, 2013). Coan (2010) integrated the emerging idea of emotional experience into a second measurement model called *the emergent variable model*. The model suggests that the synchrony of manifest indicators, which are used to measure a certain emotion, cause or rather create this emotion, which is accompanied by appraisal processes (Coan, 2010), an idea that stands in line with the assumptions of psychological construction approaches (Clore & Ortony, 2013). In contrast to the latent variable model, the synchronous response of the different parameters is determined by the elements of the situation and not by a latent brain unit or by the appraisal of the situation. Instead, appraisal gives meaning to the undifferentiated affective reaction to the situation, which results in a specific emotion. Thereby, inter- and intraindividual variability in the degree of synchrony, emotional intensity, and activated response modalities are included in the model (Coan, 2010).

The technical ability to quantify synchronous changes in multiple response variables when an emotion is induced should aid in clarifying the functionality of synchrony. This is a difficult task, due to, among other examples, the complex structure of multivariate and non-stationary indicators where distributions of variables change over time (Butler et al., 2014; Kelava et al., 2015). So far, aggregating data on an interindividual level for a time-sensitive measure such as synchrony has led to poor empirical results (e.g., Friedman et al., 2014; Hollenstein & Lanteigne, 2014; Mauss et al., 2004). Just recently, there has been a shift in the perspective on the psychological methodology that is usually applied to capture synchronous responses. Instead of looking for an aggregated effect, attention is now focused on intraindividual analyses, thus, the analysis of the course of synchronous change in one individual across a certain period of time (Bulteel et al., 2014; Kelava et al., 2015). With the present study, we want to contribute to the debate and present new empirical results regarding the synchronous response of physiological parameters in disgust. Our goal was to quantify the synchrony of physiological parameters during the emotional experience of disgust, and to find out whether the synchronous response is different for a neutral state, and how physiological synchrony is temporally related to the conscious awareness of disgust. To do so, we applied a new quantification procedure proposed by Kelava et al. (2015). The approach accounts for the complex structure of multivariate and non-stationary input variables where distributions of variables change over time.

The paper is organized as follows: First, we present empirical support for the latent variable model and basic emotion approaches, which share the perspective that emotions cause changes, as well as for the emergent variable model and psychological construction, which share the perspective that emotions emerge because of changes. Next, we provide a brief overview of the physiology of disgust. Then, we introduce the applied quantification approach and the study design. Finally, we present our results on selected peripheral physiological parameters of disgust and physiological and subjective synchrony and discuss our findings with regard to emotion theory.

Emotion causing changes

Coan (2010) described the key ideas of the latent variable model with an illustrative example: If we see the prominent bear in the woods, the latent variable model would expect the bear (or the appraisal of the bear; Lazarus, 1991; Roseman et al., 1990) to activate the fear circuit, which would result in synchronous changes (similar across different individuals) in physiological parameters (e.g., increased heart rate), behavioral level (e.g., freezing), and subjective experience (e.g., feeling afraid; see also Ekman & Cordaro, 2011; Panksepp & Watt, 2011; Tomkins, 1962). Another emotion-eliciting situation, such as a contaminated toilet would activate the disgust circuit, instead of the fear circuit, which would cause the synchronous activation of distinct response patterns of physiological (e.g., decreased heart rate), behavioral (e.g., wrinkling one's nose), and subjective parameters (e.g., feeling nauseated). Thereby, the level of activation should correspond to the intensity level of the latent emotion (e.g., Coan, 2010; Davidson, 1992; Hodgson & Rachman, 1974; Levenson, 1994; Rosenberg & Ekman, 1994).

Empirical studies that have addressed the above-mentioned hypotheses generated by the latent variable model and basic emotion approaches have produced inconsistent results. Distinct response patterns for basic emotions have been reported by some studies (e.g., Christie & Friedman, 2004; Ekman, Levenson, & Friesen, 1983; Kragel & LaBar, 2013; Kreibig, 2010; Lench et al., 2011; Levenson, 1992; Stephens et al., 2010), but such empirical evidence has been questioned by others (e.g., Barrett, 2006b; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Mauss & Robinson, 2009; Ortony & Turner, 1990; Quigley & Barrett, 2014; Russell, 2003). Further, empirical evidence for emotion-specific circuits in the brain is missing (Lindquist et al., 2012; Murphy et al., 2003; Touroutoglou et al., 2015). So far, aggregating data on an interindividual level for a time-sensitive measure such as synchrony has led to poor empirical results, meaning that synchronous responses have been weak or nonexistent (e.g., Davidson, 1978; Friedman et al., 2014; Lazarus et al., 1963; Mauss et al., 2004; Reisenzein, 2000). However, some studies have supported the idea that synchrony increases with stimulus intensity (Hsieh et al., 2011; Mauss et al., 2005). Nevertheless, studies have also identified large individual differences (Bulteel et al., 2014; Hsieh et al., 2011; Mauss et al., 2005), which were not intended by the basic emotion approach (Barrett, 2009). Therefore, interindividual variability has instead been attributed to measurement error, inappropriate methodological application, and emotion regulation (Barrett, 2006b, 2009). Due to the mixed empirical results, different theoretical perspectives, such as the emergent variable model or psychological construction, require more attention.

Emotions emerge from changes

If we go back to the prominent bear in the woods scenario, Coan (2010) described the key ideas of the emergent variable model as follows: The situation itself (i.e., a bear appears) leads to different unique problem-solving mechanisms: avoidance behavior, the automatic nervous systems prepares for quick reactions, the amygdala is activated after factoring in one's previous experience, and, finally, the collective activity results in a conscious awareness of a fear state. The conscious awareness might also occur at a later time (Clore & Ortony, 2013; Coan, 2010). In addition, psychological construction approaches place a special emphasis on physiological change (e.g., Clore & Ortony, 2013; Cunningham et al., 2013; Lindquist et al., 2013). For an emotion to emerge, a psychologically relevant situation causes or influences some sort of physiological change, which is made psychologically meaningful through continuously reinterpreting one's sensation in the light of personal goals, intentions, and values (Barrett, 2013). Therefore, physiological change, such as decreasing heart rate, does not result from the elicitation of disgust; but rather, the situation (e.g., contaminated toilet) causes the physiological change along with other reactions (e.g., wrinkling one's nose or avoidance behavior), which, through ongoing appraisal processes, results in disgust. Furthermore, individuals in the same situation - but also the same individual in the same situation at a different time - can experience different degrees of disgust, given different patterns and levels of activated indicator variables (Clore & Ortony, 2013; Coan, 2010; Lindquist et al., 2013). On the other hand, if the disgust circuit in the latent variable model or basic emotion approaches is activated on a lower level, all indicators must be activated on a lower level as well (Coan, 2010; Ekman, 1992; Tomkins, 1962).

Studies have shown that there is a great deal of heterogeneity in emotional categories between individuals (e.g., Ceulemans et al., 2012; Kuppens et al., 2007; Nezlek et al., 2008). Further, studies regarding the measurement of synchrony have emphasized large interindividual differences in synchrony during the same emotion as well (e.g., Bulteel et al., 2014; Hsieh et al., 2011), a finding that supports the propositions of the emerging theories. However, to increase the empirical foundation, Barrett (2013) called for new experimental approaches that apply a complex analysis of interacting systems over time, which we tried to establish in the present study. Furthermore, instead of looking for an aggregated effect, attention tends to focus on intraindividual analyses, that is, the analysis of the course of synchronous change in one individual across a certain period of time (Bulteel et al., 2014; Kelava et al., 2015). In addition, such new empirical approaches must address the question of temporal dynamics. The time course of physiological parameters or even physiological synchrony is largely unknown (Bulteel et al., 2014). In analyzing the synchrony of physiological parameters, most studies have induced emotions with pictures (e.g., Bulteel et al., 2014; Dan-Glauser & Gross, 2013) or film clips (e.g., Hsieh et al., 2011; Mauss et al., 2005) in a laboratory setting. In most cases, the time period from the beginning to the end of an emotional experience has been seen as equal to the duration of an emotion-evoking stimulus (Hsieh et al., 2011; Friedman et al., 2014) or different time lags such as 10 s (Mauss et al., 2005) or 8 s (Dan-Glauser & Gross, 2013) have been chosen. Nevertheless, the actual time frame of the individual's emotional experience may deviate from the presented stimulus (Bulteel et al., 2014; Waugh & Schirillo, 2012), a concept that we elaborate on in our study. Therefore, the present paper applied a new, complex, multivariate, time-sensitive, intraindividual approach in emotion research, following Barrett's (2013) appeal.

Research questions

The aim of our study was to apply advanced multivariate methods to analyze the synchrony of physiological parameters in a neutral state and a disgusting state and to determine how physiological synchrony is temporally related to the conscious awareness of disgust. Overall, we had three main research questions in the present paper: First, with *Research Question* 1, we investigated whether there was a change in physiological synchrony during the switchover ($t = 0$) to a disgusting picture and whether the change also occurred when the subsequent picture was neutral in content. Almost all of the presented approaches have proposed that some sort of synchrony between response parameters is essential for the emotional experience. However, they have differed in their allowance of inter- and intraindividual variability in the level and course of physiological synchrony. The basic emotion approaches and the latent variable model suggest that individuals will react similarly, at least at the very beginning of the emotional experience, given the biological function of emotion (Coan, 2010; Ekman, 1992; Tomkins, 1962). By contrast, the psychological construction approaches and the emergent variable model allow for different reaction patterns in different individuals and even within the same individual due to changes in appraisal processes, contextual information, personal history, learning experience, and so forth (Barrett, 2006a; Clore & Ortony, 2013; Coan, 2010).

Second, with *Research Question* 2, we took a closer look at the temporal course of physiological synchrony. In particular, we were interested in identifying when synchrony would increase, reach its maximum, and decrease after a picture change. Due to the quantification method we applied and therefore due to the high temporal resolution that resulted, we were able to observe real-time changes in physiological synchrony without aggregating data across a longer time period. The basic emotion approaches and latent variable model suggest that the stimulus (or the appraisal of the stimulus) activates some form of disgust circuit in the brain, which then affects physiological, behavioral, and subjective response parameters (Coan, 2010; Ekman & Cordaro, 2011; Panksepp & Watt, 2011). On the basis of this idea, physiological synchrony should start to increase after the disgust circuit causes the change, hence, after the stimulus presentation. By contrast, the psychological construction approaches and the emergent variable model see physiological change and the determinants of the situation as evolving into an emotional experience, without defining an explicit trigger (Barrett, 2006a; Clore & Ortony, 2013; Coan, 2010).

With *Research Question* 3, we explored temporal differences in the time course of the estimated physiological synchrony and the continuously reported subjective disgust ratings of the pictures. We would like to emphasize the exploratory character of this research question because, due to the aggregation of data across time (e.g., Friedman et al., 2014; Mauss et al., 2005) or the use of retrospective subjective ratings with questionnaires (e.g., Friedman et al., 2014; H. S. Schaefer et al., 2014), direct comparisons of temporal courses have not been performed in the past. To the best of our knowledge, we are the first to report such a comparison. According to the basic emotion approaches and the latent variable model, all manifest indicator (physiological and subjective) variables would be expected to change more or less synchronously (Coan, 2010; Ekman, 1992). By contrast, according to the psychological construction approaches and the emergent variable model, the conscious awareness of an emotional state evolves through the joint activity of response modalities, thereby allowing conscious awareness to appear at a later time (Clore & Ortony, 2013; Coan, 2010).

5.2. Method

Participants

In total, 51 participants were invited to take part in our study. In our screening questionnaire, we collected data on chronic illnesses, mental illnesses, psychotherapeutic treatment, drug ingestion, smoking, alcohol consumption, substance abuse, vision, and pregnancy. Four participants were excluded from the study due to mental illnesses, drug ingestion, and chronic illness. Five additional participants were excluded because they had smoked or consumed caffeine 2 hr before the experiment. The final sample consisted of 42 participants (50% female). The participants' mean age was 30.29 years (SD = 12.19, age range: 21-67 years). In total, 60% of the sample were students (52% psychology, 48% other), and 40% were working or retired. All participants took part in a drawing for three 25-Euro Amazon gift cards. The student participants additionally received course credit. The experiment was approved by the local ethics committee.

Procedure

The participants were welcomed and seated on a normal office chair in a 4 × 6 m laboratory room at a distance of 0.5 m from the computer screen. First, the experimenter informed the participants about the procedure. Participants were told that, during the experiment, they were going to see pictures that could provoke feelings of disgust. They were asked to continuously report their feelings of disgust by turning a rating dial while looking at a picture. Zero on the rating dial meant *not disgusting at all*, whereas 10 meant *very disgusting*. After each picture, they were asked to turn the dial back to 0 and rate the following stimuli, beginning at 0. Furthermore, they were supposed to keep in mind that they were to rate only their feelings of disgust. They were not supposed to rate feelings such as happiness or sadness or to report whether they liked or disliked a picture. At the end of the introduction, the participants had to fill out an informed consent form and a screening questionnaire; afterwards, the psychophysiological sensors were attached to their bodies. Then a sample baseline measurement of 5 min was taken. Throughout this phase, participants were asked to stay still, but there was no neutral material presented to them.

The experiment was divided into two blocks. First, the participants viewed a baseline block with 21 neutral pictures from the International Affective Picture System (IAPS; Lang et al., 2008) followed by a disgust block, again with 21 pictures (15 disgust and 6 neutral). Within each block, pictures were displayed for 7 s, with a black screen presented for 3 s between pictures. The duration of time for which the pictures were displayed was similar to ones that have been used before (e.g., Lang et al., 1993; Stark et al., 2005). Within each block, all pictures were shown in a random order. Each block lasted for 3 min 48 s. After each block, participants filled out questionnaires. For the programming of the sequence of the pictures as well as for the processing of the experiment, we used the software Matlab R2011b (MathWorks, 2011).

Self-report measures

To control whether we evoked only feelings of disgust or other negative emotions, participants were asked to fill out the *Differential Affect Scale* (DAS; Merten & Krause, 1993). The DAS is a German translation of the *Differential Emotions Scale* (DES; Izard, Dougherty, Bloxom, & Kotsch, 1974), which has previously been applied to differentiate between disgust and fear (e.g., Marzillier & Davey, 2005). The DAS assesses 10 emotions such as amusement, surprise, fear, disgust, anger, and so forth. Each emotion is measured with three adjectives (e.g., *sick, nauseated*, and *disgusted* for *disgust*) on a 5-point Likert scale (0 = *not at all*, 5 = *very strong*). Participants had to rate retrospectively how strongly they felt a particular emotional adjective while viewing the pictures. The DAS was administered twice, one time after the neutral pictures and one time after the disgusting pictures. Thus, it reflects the participant's overall rating of the neutral baseline block and overall rating of the disgusting block. Due to the applied measurement setup, the rating dial response, which constituted the second form of self-report measure in this study, is described in the following paragraph.

Physiological measures

The electrocardiogram, respiration, and skin conductance responses as well as the rating dial were measured with the Biopac MP 150 System and the AcqKnowledge 4.2 Software (Biopac Systems Inc, 2011). The data were sampled with a sampling frequency of 250 Hz. *Electrocardiogram*: To record the electrocardiogram, we used pregelled Ag/AgCl electrodes. The white electrode was placed at the beginning and the black electrode at the end of the sternum. The red electrode was attached to the left side of the chest, on the third rib, bottom up. The gain was set to 2000 using a normal ECG acquisition mode. *Skin conductance response*: To record electrodermal activity, two Ag/AgCl electrodes (diameter, 8 mm) filled with isotonic gel (0.5% saline in a neutral base) were attached to the palm of the nondominant hand. The GSR100c used a constant voltage of 0.5 V. *Respiration*: To record respiratory effort, the TSD201 transducer belt was attached around the participant's thorax. The participant was told to exhale completely, then the belt was tightened when the circumference was at its minimum. The respiration transducer was attached to the participant in a sitting position, which the participant stayed in throughout the experiment. The signals were filtered by a 10-Hz low-pass filter. *Continuous rating dial*: In each block, participants continuously rated their personal feelings of disgust by turning a rating dial. The online rating method does not appear to disturb participants' emotional experiences (Mauss et al., 2005). The potentiometer was attached to an UIM100C module from the biopac system and sampled at a sampling frequency of 250 Hz. The potentiometer has a scale ranging from 0 (*not disgusting at all*) to 10 (*very disgusting*) and an incoming voltage range from 0 to 5 V. The dial was placed in a hemicycle on a 20 × 13 × 8 cm box. Before each block, participants saw a written introduction with explanations about the dial on the computer screen.

Material

Disgust has been elicited successfully by using pictures in the past (e.g., Lang et al., 1993; Overbeek, van Boxtel, & Westerink, 2012; Ritz, Thons, Fahrenkrug, & Dahme, 2005; Rohrmann & Hopp, 2008; Schienle, Stark, & Vaitl, 2001; Stark et al., 2005). We displayed 21 neutral IAPS pictures[7] in the baseline block, followed by 15 disgust and 6 neutral pictures[8] in the disgust block. The pictures were part of a larger study on disgust in general; therefore, we randomized the order of the pictures to avoid sequence effects (e.g., Rohrmann & Hopp, 2008; Stark et al., 2005). Neutral pictures were included in the disgust block because we wanted to lighten the, in some parts, heavy content of the pictures. We chose the content of

[7]IAPS picture numbers in the baseline block: IAPS no. = 5201, 5220, 5260, 5551, 5600, 5635, 5661, 5720, 5740, 5891, 7002, 7006, 7009, 7010, 7030, 7039, 7050, 7052, 7150, 7224, 7235.

[8]IAPS picture numbers in the disgust block: IAPS no. = 1270, 3100, 3130, 3150, 5660, 5780, 7004, 7025, 7211, 7175, 7359, 9290, 9301, 9320.

the pictures according to the five disgust-eliciting domains presented by Schienle, Walter, Stark, and Vaitl (2002): death, body secretions, hygiene, spoilage, and oral rejection. We picked three content-related pictures from each disgust category. In order to complete the pictures within each category, we selected seven pictures (e.g., a toilet with excrement, a picture of puke in a car) that did not come from the IAPS data base. To ensure that our selected stimuli evoked disgust, we conducted a prestudy (N = 10). Table A1 in the Appendix displays the mean rating dial scores for the five most disgusting pictures. The pictures that belonged to Schienle et al.'s (2002) disgust categories of death, hygiene, and oral rejection were rated as the most disgusting. These results were confirmed in our main study. Even though participants viewed all 21 pictures during the disgust block, we chose only the five most disgusting pictures for further analysis because research has repeatedly emphasized that a lack of emotional intensity might be the reason for the lack of empirical support (Hollenstein & Lanteigne, 2014; H. S. Schaefer et al., 2014). One goal of this paper was to evaluate the effect of emotional stimuli on synchrony. Therefore, we wanted to ensure that disgust was evoked.

Applied quantification of synchrony

For the quantification of synchrony, we applied a new time-frequency-based latent variable approach by Kelava et al. (2015). We briefly introduce the key ideas of the approach below. First, we explain the indicator variables of the latent synchrony variable, and then we describe the approach of state-space models. The interested reader can consult Kelava et al. (2015) for a more detailed description.

Indicator variables

Many of the most powerful and effective signal processing algorithms rely on the assumption of stationarity. Signals of physiological origin, however, often do not comply with this assumption. As a first step, to derive the indicator variables, we therefore resort to time-frequency analysis, which is able to deal with non-stationary signals. The time-frequency domain jointly displays the evolution of a signal's power in time and frequency, and has been regularly used in the signal processing community to analyze physiological data (Boashash, 2015; Hlawatsch & Auger, 2013; Lander & Berbari, 1997; Novak & Novak, 1993; Peláez-Coca, Orini, Lázaro, Bailón, & Gil, 2013). The pairwise coherence measure for each time instant and frequency point is defined by:

$$C_{XY}(t,f) = \frac{S_{XY}(t,f)}{\sqrt{S_{XX}(t,f)S_{YY}(t,f)}}, \tag{5.1}$$

where $S_{XY}(t,f)$, $S_{XX}(t,f)$, and $S_{YY}(t,f)$ are the cross and auto time-frequency distributions of the time signals $x(t)$ and $y(t)$, respectively.

In this paper, as in Kelava et al. (2015), the time-frequency coherence was smoothed by performing a two dimensional filtering with a Gaussian kernel, to ensure that $C_{XY}(t,f)$ has the below mentioned properties. The parameters of the Gaussian kernel define the time and frequency resolution capability of $C_{XY}(t,f)$ and were chosen on the basis of synthetic test signals (see Kelava et al., 2015), such that both sharp changes in time (time resolution) and individual frequency regions of interest (frequency resolution) can be resolved. Time-frequency coherence displays the linear coupling between $x(t)$ and $y(t)$ at time instant t in frequency f. It is bounded by 0 and 1. If $x(t)$ and $y(t)$ are completely uncorrelated, their coherence equals zero. If, on the other hand, $|C_{XY}(t,f)|^2 = 1$, $y(t)$ can be fully predicted at time instant t from $x(t)$ by a linear system. When $|0 < C_{XY}(f)|^2 < 1$, as described in Bendat and Piersol (1990), this can be due to, for example, noise entering the measurements or because of a nonlinear functional relation between the signals. The indicator variables of the latent synchrony variable at time t can then be obtained by averaging $C_{XY}(t,f)$ at time t across all frequencies within the detected delineated frequency region. Our approach is thus based on time-varying frequency specific coherence analysis followed by state-space modeling. There is also an ongoing and interesting research on cross-frequency phase coupling (mainly applied to EEG data). As discussed, for example, in Cohen (2008) the cross-frequency phase coupling approach has both advantages and limitations.

The power of the considered physiological signals is not evenly spread out in the time-frequency plane but is instead concentrated in delineated regions (Kelava et al., 2015; Muma et al., 2010). In the case of the ECG signal, most of the power is concentrated in the pulse frequency region that can vary in a broad frequency range approximately between 0.67 and 3.33 Hz (corresponding to 40 - 200 beats per minute) and its harmonics, which are integer multiples of the pulse region. The power of the respiratory signal, on the other hand, is mostly concentrated in a single region that represents the respiratory cycles. The EDA signal is somewhat different because, in general, it does not have a cyclical nature but rather trends and abrupt changes. Most of its signal power is concentrated in the lowest frequencies of the time-frequency plane (for a detailed and thorough analysis of EDA signals, see Lim et al., 1997). In this paper, we have used an algorithm that was published in Kelava et al. (2015) and Muma et al. (2010) to detect for each time-instant the above described delineated regions. The interested reader is referred to these publications, as well as to additional material[9], where we provide an annotated tutorial-style Matlab code that exactly describes how the time-frequency coherences and the delineated regions were computed.

By detecting the above-described delineated regions and applying the concept of time-frequency coherence, as defined in Equation 1, we were able to determine pairwise, scalar, time-varying coherence measures between two signals. In this paper, we considered three physiological signals, each of which had one particular region of interest. Normally, this would result in six indicators. For example, one indicator is the average coherence of respiration and ECG within the respiration region for each time instant. The intercorrelation between these two cardiovascular signals is well known and has been analyzed in time-frequency (and time-scale) domain (Keissar, Davrath, & Akselrod, 2009). However, in a different way from Kelava et al. (2015), we excluded the coherence of the EDA and ECG signals in the pulse frequency range, even though we could observe small signal powers in the pulse frequency range, which are correlated to the ECG signal. The indicator was excluded because these high frequency EDA components are in fact interferences from the ECG signal due to the measurement process rather than physiologically generated signal content (Boucsein, 1988). The remaining five indicator variables were used to determine the overall synchrony, as detailed in the following section. Figure 5.1 exemplifies the five indicators for the duration of the viewing of a neutral picture followed by a disgusting picture for a selected Participant A. At $t = 10$ s, the picture switched from neutral to disgusting.

State-space models

As a second step, a latent state-space modeling approach was introduced by Kelava et al. (2015) to specify an overall synchrony measure that was based on the coherence indicators that were described in the previous section. The basic idea of the state-space procedure is to specify a measurement model that operationalizes a

[9] Additional material: `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4299432/bin/DataSheet1.ZIP`.

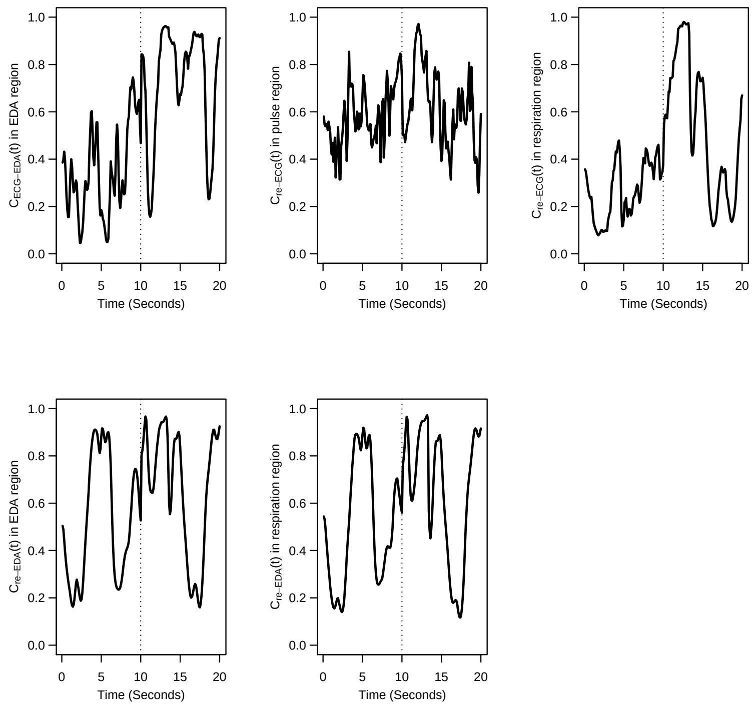


Figure 5.1. An example (Participant A) of the five scalar time-varying coherence measures that were used as indicators of the latent synchrony variable. At $t = 10$ s, the picture switched from neutral to disgusting (highlighted by the dashed vertical line). $C_{ECG-EDA}(t), C_{re-ECG}(t)$, and $C_{re-EDA}(t)$ refer to the pairwise time-varying coherences in the ECG, respiration, and EDA signals. The time-frequency regions of interest for the coherence analysis were the pulse, respiration, and EDA regions. The scalar coherence value that is plotted represents the average pairwise coherence between two signals, whereby the average was taken for each time point over the frequencies contained in a given region of interest.

latent variable that represents an underlying state of the overall system-wise synchrony. This measurement model (see Equation 2) describes the relationship between the manifest coherence indicators from above and a latent variable (ξ), which represents the overall synchrony for a specific time point (i.e., a state variable of synchrony). Given a finite number of measured psychophysiological signals, the overall synchrony measure is not an entity that refers to all physiological changes in the body but only to those that were chosen in the study. In addition to the measurement model, a structural model (see Equation 3) describes the regression of a state of synchrony relative to its previous states.

For a given individual i, the measurement and structural models are:

$$z_{it} = H_t \xi_{it} + R_t e_{it} \tag{5.2}$$

and

$$\xi_{it} = F_t \xi_{i(t-1)} + G_t u_{it} + Q_t \eta_{it}, \tag{5.3}$$

respectively. In the measurement model (Equation 5.2), a given p-dimensional observed variable z_{it}, at time point t, is regressed onto a q-dimensional (latent) state vector ξ_{it}, where e_{it} is a p-dimensional residual vector (white noise). The variable z_{it} hereby represents the above-described pairwise coherence indicators, and ξ_{it} denotes the individual's time-varying overall synchrony that is sought for an individual i at time point t. H_t and R_t are time-varying coefficient matrices of dimensions $p \times q$ and $p \times p$ that describe the time-varying relations between z_{it} and ξ_{it}, and z_{it} and e_{it}, respectively.

In the structural model (Equation 5.3), the state vector ξ_{it} is (auto-)regressed on a previous state vector $\xi_{i(t-1)}$ and on an m-dimensional covariate vector u_{it} (e.g., an intervention), where η_{it} is a q-dimensional latent residual (white noise) vector. Again, F, G, and Q are time-varying coefficient matrices of dimensions $q \times q$, $q \times m$, and $q \times q$, respectively. The coefficient matrix H_t displays the time-dependent strength of the relations between the overall synchrony and its indicators. In the present study, the analysis was conducted with the R-package *dse* by P. Gilbert (2013). The coherence indicators were calculated with Matlab, and the R software (R Core Team, 2012) was used for the state-space model. Details and codes are provided in Kelava et al. (2015).

Data analysis

Manipulation check

Subjective data. By using the DAS (Merten & Krause, 1993), we aimed to ensure that only the emotion of disgust and not fear, surprise, anger, or sadness would be evoked in our experiment. To evaluate this, in a first step, we calculated the $DAS_{emotion}$ - $DAS_{baseline}$ difference for each emotion and each participant. In a second step, we computed the mean of the differences (M_D) for each emotion. In order to test whether disgust was evoked, we compared the obtained means of differences by using two-tailed paired t tests. Because we were interested only in whether disgust was induced in a larger quantity than fear, surprise, anger, or sadness, we compared the M_D of disgust against that of fear, surprise, and so forth. We applied the*Bonferroni* correction to control the familywise error rate. We calculated Cohen's d as a measure of effect size.

Physiological features. In addition to our manipulation check of the subjective data, we also evaluated whether we were able to evoke a measurable effect of disgust in the physiological signals. For this analysis, we selected Pictures 9 to 13 from the neutral baseline block and the five most disgusting pictures identified in the prestudy. We chose the middle pictures from the baseline block for two reasons: first, to assure that the initial excitement had already decreased. We believed that after participants were confronted with eight neutral baseline pictures, these effects would be reduced. Second, we expected excitement in the later baseline pictures, induced by the expectation that the upcoming pictures would induce disgust. The pictures were displayed in a random order. We calculated HR, HRV, SCR, and ID. Our evaluation of these physiological features was independent of our evaluation of the synchrony measure and served only as parameters for conducting our manipulation check. The HR, HRV, SCR, and ID values were estimated as follows. *HR*: After data acquisition, a low-pass filter was applied with a cutoff frequency of 20 Hz. Motion artifacts in the ECG signals were removed by applying a method proposed by Strasser et al. (2012). The HR signal in beats per minute was estimated with a peak detection algorithm (see also Strasser et al., 2012). *HRV*: We used the time difference between two detected R-peaks to compute the HRV estimate. To extract the R-R intervals, we applied the QRS detector by Pan and Tompkins (1985), as implemented by Clifford (2002). *SCR*: We used the largest increase in skin conductance that occurred between 0 s and 4 s after the picture onset as the

SCR estimate (e.g., Lang et al., 1993; Stark et al., 2005). *ID*: The endpoints of the inhalation and exhalation process represented the maxima and minima, respectively, in the respiration signal. We determined these via a peak-finding algorithm. The differences between consecutive maxima and minima served as estimates of ID. All physiological features were downsampled from 250 Hz to 10 Hz to reduce the size of the data.

Main analysis

To examine the change in physiological synchrony that occurred during the change in pictures, we again chose only the five middle baseline pictures and the five most disgusting pictures as determined by the prestudy. To examine the effect of a picture, we focused our analysis on a time interval that contained a picture change. The interval included the 7 s in which a new picture was displayed as well as the 3 s before and after, when a black screen was shown. For this time interval, we measured the subjective experience of disgust via the rating dial and computed the overall physiological synchrony ξ_{it} as detailed above. To establish an average time course of the overall physiological synchrony for the neutral and disgusting pictures, we aggregated ξ_{it} across all individuals i (maintaining the time course) of each class separately. In this way, an average time course of physiological synchrony was obtained for the neutral and the disgusting pictures (see Figure 5.4). We performed the same type of aggregation across individuals for the rating dial values, resulting in an average time course for the neutral and disgusting pictures. Because we were interested in only the reactions to the pictures and not in the absolute level of synchrony, we subtracted the sample mean from the average overall synchrony in order to center the data.

5.3. Results

Manipulation check

Differential affect scale

The means, standard errors, t values, corrected p-values, and Cohen's d are presented in Table A2 in the Appendix. Due to our manipulation, disgust had a significantly larger DAS value compared with fear, surprise, anger, or sadness. Figure 5.2 displays the mean DAS scores for disgust, fear, surprise, anger, and sadness after the neutral baseline block and after the disgust block. Disgust had the statistically significantly largest increase compared with the other negative emotions. On the basis of these results, we were confident that we evoked primarily the emotion of disgust.

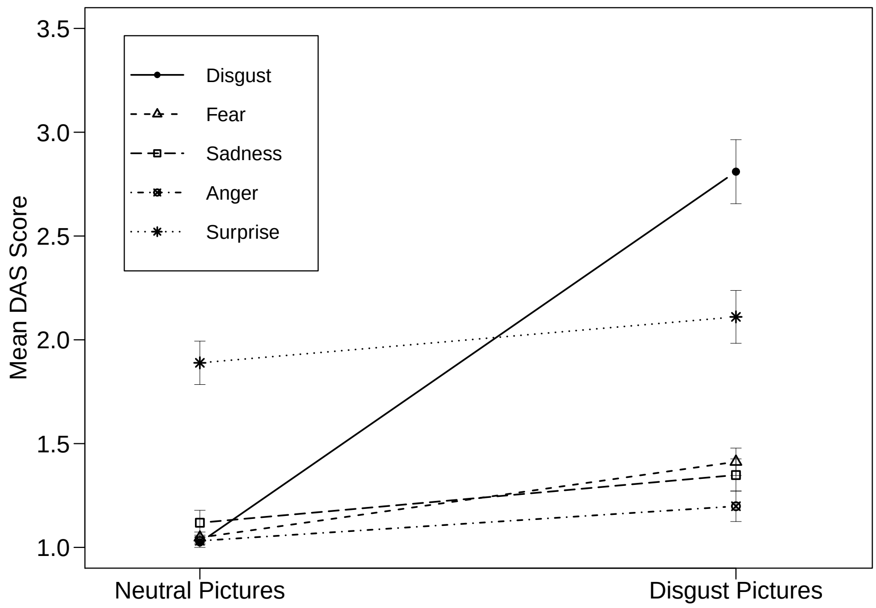


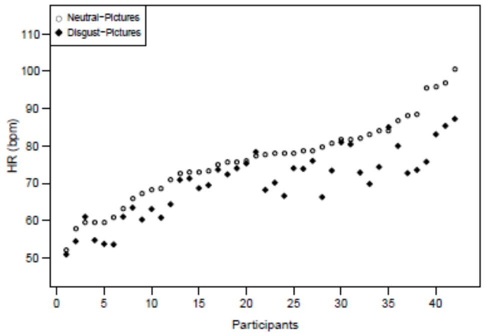
Figure 5.2. Manipulation check of the mean scores from the Differential Affect Scale (DAS; Merten & Krause, 1993) for neutral and disgusting pictures.

Physiological data

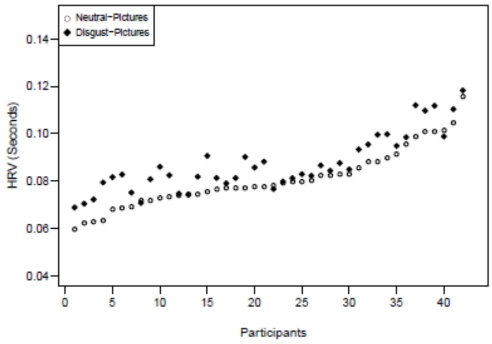
We estimated the mean HR, HRV, SCR, and ID signals for the neutral baseline and disgust blocks (see Figure 5.3). While participants viewed the disgusting pictures, HR[10] decreased, whereas HRV and SCR increased. ID did not differ between the neutral and disgusting pictures. The differences between the neutral baseline block and the disgust block (disgust t_2 - neutral t_1) was shown to be significant for HR, HRV, and SCR (see Table A3 in the Appendix). Thus, our signal measures confirmed that the pictures evoked sympathetic and parasympathetic activation and physiological reactions to disgust. In addition to the subjective data, the

[10] The purpose of Figure 5.3 was to display the intraindividual differences between the HR when showing neutral pictures, compared to the HR when showing disgusting pictures. The absolute value of the HR was not of particular importance.

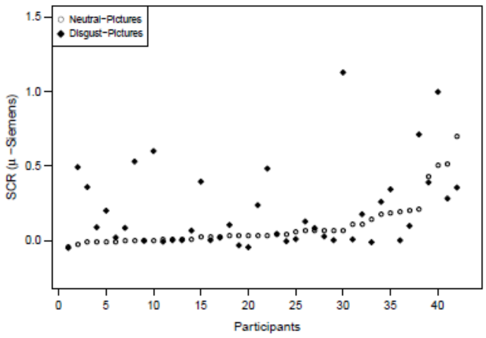
physiological data provided further evidence that disgust, as opposed to any other negative emotion, was elicited by our stimulus material.



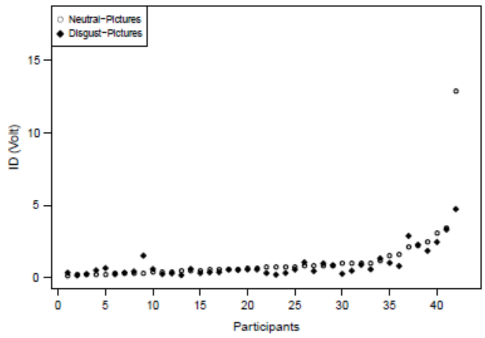
(a) Mean heart rate (HR) in beats per minute (bpm)



(b) Mean heart rate variability (HRV) in seconds



(c) Largest increase in skin conductance response (SCR) between $t = 0$ and $t = 4$ s in micro Siemens (μ-Siemens)



(d) Inspiratory depth (ID) in Volts

Figure 5.3. Mean physiological parameters for the neutral baseline and disgusting pictures. Participants are plotted in ascending order of the neutral pictures.

Research Question 1

With our first research question, we explored whether there was a change in physiological synchrony when the picture switched to a disgusting picture and whether the change was different than when the picture came from the neutral picture set. First, we want to highlight that there were large interindividual differences in physiological synchrony. Figure A1 in the Appendix shows the average physiological synchrony (not centered) and average rating dial values for the neutral and the disgusting picture changes for each participant, respectively. Participants differed in their physiological synchrony response in terms of level, time course, beginning, and end. Acknowledging the interindividual differences, we proceeded with Research Question 1 by investigating the average time course of the participants' reactions across all participants. Clearly, the average should convey similarities, whereas the standard deviation should express variability. We also considered other statistical characteristics of the population such as the median and the median absolute deviation, both of which are less sensitive to outliers than the sample mean and standard deviation (Zoubir, Koivunen, Chakhchoukh, & Muma, 2012). Because the results were very similar to the results for the sample mean and standard deviation, only the classical measures are reported in this paper. Because the nonrobust and robust measures showed similar results, we were also able to conclude that the empirical probability distributions of the time courses were not heavy tailed (i.e., did not contain outliers; Tukey, 1979).

Figure 5.4 shows the average time course of physiological synchrony (ξ_{it}) computed over all individuals i for the neutral and disgusting pictures. The x-axis of Figure 5.4 displays the times when the pictures changed. The time t = 0 s corresponds to the moment when a new picture (either neutral or disgusting) was shown to the participant. This picture was shown for 7 s. Before and after the picture, a black screen appeared for a duration of 3 s (i.e., a black screen appeared at t = -3 s and at t = 7 s). Therefore, t < -3 s and t > 7 s correspond to the previous and subsequent pictures, respectively. The y-axis of Figure 5.4 depicts the average time course of the overall physiological synchrony for the neutral and disgusting pictures. With Research Question 1, because we were mainly interested in the changes in the reactions to the pictures, the sample mean was subtracted from the average overall synchrony in order to center the data. Figure 5.4 also shows the average time course of the rating dial results, which captured participants' subjective feelings of disgust. Again, the average was computed over all individuals in each class separately. For visual clarity only, the time course of the rating dial results was scaled by a constant factor so that the subjective and physiological signals could be better displayed in a single plot. At time point t = 0, at the moment that the picture changed, a significant difference in the level of synchrony between neutral and disgusting pictures was observed ($M_{Disgust}$ = 0.12; $M_{Neutral}$ = -0.04), $t(41)$ = -15.12, p < .01. At time point t = -2 s, on the other hand, when participants viewed the black screen without knowing whether a neutral or a disgusting picture would appear, the level of synchrony did not differ significantly between the neutral and disgusting picture classes ($M_{Disgust}$ = -0.02; $M_{Neutral}$ =

-0.01), $t(41) = -0.42$, $p > .67$. The differences between the average synchrony for the neutral and disgusting picture classes are highlighted in Figure 5.5, which plots the difference between the two time courses by subtracting the neutral from the disgust time course. Thus, a positive value in Figure 5.5 means that, at this point in time, physiological synchrony as a reaction to disgusting pictures was higher than physiological synchrony as a reaction to neutral pictures. The difference had its maximal value at the moment at which the pictures changed. To evaluate the variability in the synchrony time courses, Figure 5.6 depicts the standard deviation of synchrony for all participants for each time point during the neutral and disgusting picture sequences. The x-axis of Figure 5.6 is identical to the previous figures. It is interesting that the smallest variability in the participants' synchrony occurred at the moment when synchrony was highest. In other words, this means that at this time point, the participants' reactions were most similar.

Research Question 2

The second research question concentrated on the temporal course of physiological synchrony when disgusting and neutral pictures were displayed. In particular, we were interested in determining the point in time (with reference to the time of the appearance of a new picture) when synchrony increased, reached its maximum, and decreased. We would like to emphasize here that, to the best of our knowledge, the temporal description of a physiological synchrony response in disgust and nonemotional contexts constitutes a novelty in itself. First, we looked for the beginning of the slope, representing the increase in synchrony. This is defined by the closest zero crossing of the first derivative, and this corresponds to the first sample for which the first derivative is positive. For the average synchrony for disgusting pictures, this occurred at $t = -1.2$ s (i.e., while the black screen was presented before the emotional stimulus; see Figure 5.4). Thus, on average, physiological synchrony began to increase while participants viewed the black screen, without the influence of an emotional stimulus. The maximum level of physiological synchrony was reached 0.3 s after the disgusting pictures appeared. On average, synchrony then decreased continuously until the first derivative crossed the zero line at $t = 2.4$ s after the picture was shown. Thus, when viewing these time points as the starting and ending points of the synchronization for disgust, the reaction lasted for an average of 3.6 s. We next analyzed the average time course of synchrony when neutral pictures were shown. Similar to the disgusting pictures, the beginning of the slope repre-

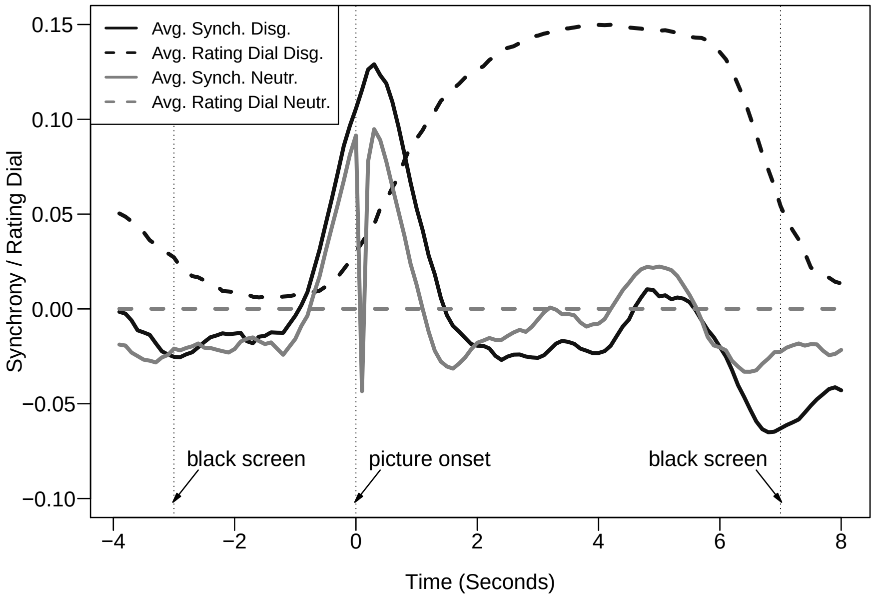


Figure 5.4. Average (Avg.) time course of the overall physiological synchrony (Synch.) for the neutral (Neutr.) and the five most disgusting (Disg.) pictures. Synchrony was averaged across all individuals for the neutral and disgusting pictures separately. The same type of averaging across individuals was performed for the rating dial values, resulting in two average time courses. The absolute level of synchrony was subtracted by centering the data. The x-axis represents the time in seconds. The point in time $t = 0$ s corresponds to the moment when a new picture (either neutral or disgusting) was shown to the participant. This picture was shown for 7 s. Before and after the picture, a black screen appeared for a duration of 3 s (i.e., a black screen appeared at t = -3 s and at t = 7 s), $t <$ -3 s and $t >$ 7 s correspond to the previous and subsequent pictures, respectively. At the moment when the picture changed ($t = 0$ s), on average, synchrony was significantly higher for the disgusting than for the neutral pictures.

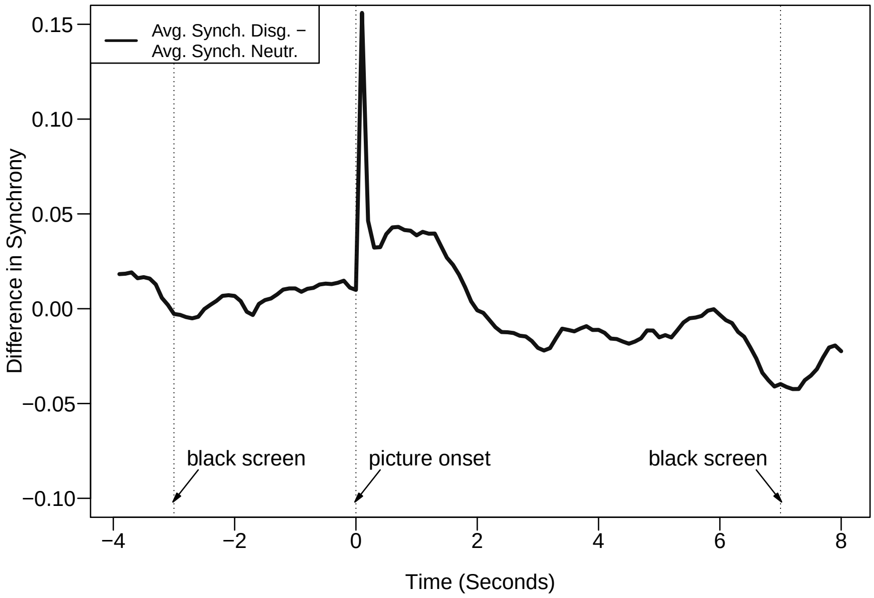


Figure 5.5. Difference between the average (Avg.) time course of the overall physiological synchrony (Synch.) for the neutral (Neutr.) and the five most disgusting (Disg.) pictures. The x-axis is the same as in Figure 5.4. A positive value on the y-axis means that physiological synchrony as a response to the disgusting pictures was higher than physiological synchrony as a response to the neutral pictures. The difference was highest at the moment at which the picture changed.

senting the increase in synchrony occurred on average at t = -1.3 s before a new picture was shown (i.e., while participants viewed the black screen that preceded the picture). At the moment when a neutral picture appeared, synchrony showed a marked decrease for a brief duration of $t = 0.1$ s. Then, at $t = 0.2$ s, synchrony reached approximately the value it had before the picture had appeared (at t = -0.1 s) before continuously decreasing until the first derivative crossed the zero line from the bottom for the first time at $t = 1.3$ s. Again, taking the starting and ending points of the slope as the beginning and end of the synchronization reaction, for the neutral pictures, the reaction lasted an average of 2.9 s. In particular, compared with the disgusting pictures, the end point of the slope occurred 1.1 s earlier. All temporal characteristics that we analyzed are summarized in Table 5.1.

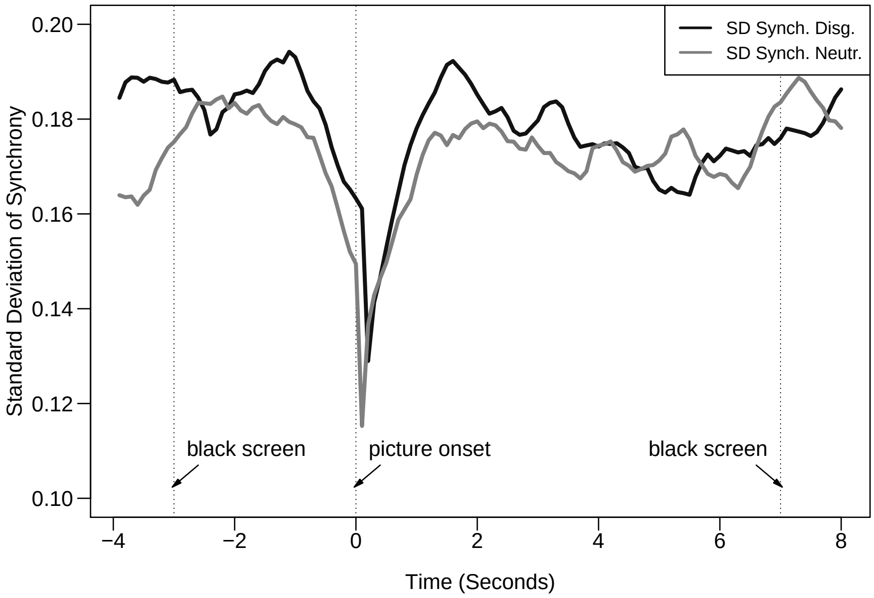


Figure 5.6. Standard deviations (SD) of the physiological synchrony (Synch.) of all participants for the five most disgusting (Disg.)- and the five middle neutral (Neutr.) baseline picture sequences. The x-axis is the same as in Figures 5.4 and 5.5. The variability between participants was smallest in the interval capturing the moment at which a new picture appeared.

Research Question 3

With our third research question, we were interested in the differences in the time courses of physiological synchrony and the subjective ratings. First, we would like to point out that participants turned the rating dial when they viewed the disgusting pictures, but they did not turn it when they viewed the neutral pictures. This is a further indication of a successful emotion manipulation. Our results showed that, on average, subjectively reported disgust feelings increased much more slowly than the physiological synchronization (see Figure 5.4). Whereas, on average, physiological synchrony reached its maximum 0.3 s after the picture change, the rating dial reached its maximum at $t = 3.9$ s. Also, the subjective disgust ratings expressed with the dial lasted longer (on average 9.8 s) than the physiological reaction (see Table 5.1).

Table 5.1.

Time Values of Interest in Describing Change in the Average and Centered Physiological Synchrony and Average Subjective Ratings of Disgusting or Neutral Pictures According to Figure 5.4

	Disgust			Neutral	
	$t_{Synch.}$	Synch.	t_{RD}	$t_{Synch.}$	Synch.
Beginning of increase	-1.2 s	-0.02	-1.3 s	-1.3 s	-0.03
1. Maximum	0.3 s	0.14	3.9 s	-0.1 s	0.10
2. Maximum	-	-	-	0.2 s	0.11
End of decrease	2.4 s	-0.03	8.5 s	1.6 s	0.02

Note. t = time values of interest; Synch. = overall physiological synchrony averaged across all participants and centered by subtracting the sample mean; RD = average rating dial.

5.4. Discussion

The quantification of the synchrony of physiological reactions to emotional stimuli has recently attracted much attention. However, to our knowledge, the synchrony of selected physiological parameters in disgust had yet to be investigated. Thus, we explored short-term intraindividual changes in physiological synchrony and subjective ratings of disgust and compared them with a neutral stimulus in terms of level, beginning, duration, and end. In the following, we will discuss the results in more detail and embed them into the current empirical findings and theoretical approaches.

With our first research question, we aimed to explore changes in physiological synchrony during the onset of disgusting and neutral pictures. The results showed that, on average, heart activity, respiration, and skin conductance were more tightly coupled at the instant at which a disgusting picture was presented, compared with a neutral picture. This observation is consistent with assumptions of the basic emotion approaches and the psychological construction approaches, as they postulate some form of synchrony between response parameters during an emotional episode (Clore & Ortony, 2013; Ekman, 1992; Levenson, 1994). However, the basic emotion approaches and the latent variable model expect the changes to be similar for different individuals (Coan, 2010; Ekman, 1992), at least at the very beginning of the emotional experience and when the stimulus is intense (e.g., Ekman & Cordaro, 2011). In our study, while viewing disgusting pictures, participants reacted most similarly when synchrony was at a maximum, less than 1 s after the disgusting picture was presented, which coincides with what is posited by the basic emotion approaches (Ekman & Cordaro, 2011; Izard, 2011; Levenson, 2011). On the other hand, our data supports the presence of large inter- and even intraindividual differences in the course and level of physiological synchrony. The results confirm previous studies that reported large interindividual differences in response synchrony (Bulteel et al., 2014; Hsieh et al., 2011; Mauss et al., 2005). On average, as described above, the participants reacted with an additional increase in synchrony when the following picture was disgusting and with a short-term decrease when the following picture was neutral. However, some participants' reactions were completely different; for example, they experienced a decrease or delayed increase in synchrony when viewing disgusting pictures. Further, the individual reaction times appeared to vary between individuals and also seemed to be different for different pictures. Keeping the basic emotion assumptions in mind, emotion regulation might be at least a partial explanation for the interindividual differences. The influence of emotion regulation strategies (e.g., suppression or reappraisal) on the behavioral, subjective, or physiological levels of disgust has been investigated by numerous empirical studies (e.g., Demaree et al., 2006; Gross, 1998; Gross & Levenson, 1993; Rohrmann, Hopp, Schienle, & Hodapp, 2009). Further, the influence of emotion regulation on response synchrony has been shown in the past as well (Butler et al., 2014; Dan-Glauser & Gross, 2013). Despite the fact that we did not give any explicit emotion regulation instructions in our study, we cannot rule out the possibility that some participants regulated their emotions automatically, without explicitly being told to do so (e.g., Gyurak, Gross, & Etkin, 2011). Besides emotion regulation, as one possible source for the high interindividual variability in physiological synchrony, our sample had a large range and high standard deviation in age. Previous studies have usually collected data on female participants (Bulteel et al., 2014; Butler et al., 2014; Gentsch et al., 2014; Mauss et al., 2005; H. S. Schaefer et al., 2014). Even though these previous participants had a lower standard deviation for age and the same gender, high interindividual differences in physiological response synchrony were reported (Hsieh et al., 2011; Mauss et al., 2005), which speaks against a strong influence of age on our results. However, the influence of age on physiological synchrony during emotional experience is an interesting question that should be addressed in the future. The above given explanations for interindividual differences, such as emotion regulation, age, or gender are indeed plausible but they cannot account for the intraindividual differences in our study. Further, according to the basic emotion approaches, these strong inter- and intraindividual variations in synchrony should not really be expected (Barrett, 2009). By contrast, the psychological construction approaches and the emergent variable model allow for inter- and intraindividual variability in response synchrony due to the different contexts (e.g., picture content), current circumstances, learning experiences, and life history (Barrett, 2009; Clore & Ortony, 2013; Coan, 2010; Quigley & Barrett, 2014). Summing up, physiological synchrony was significantly higher at the onset of a disgusting picture, than at the onset of a neutral picture. This empirical finding is in accordance with the propositions of the basic emotion approaches and the latent variable model, but at the same time, this result does not contradict the physiological construction approaches or the emergent variable model either. However, the great interindividual and intraindividual differences in our results tend to support the key ideas of the psychological construction approaches and the emergent variable model.

With our second and third research question, we focused in particular on the temporal course of physiological synchrony and the conscious awareness of the emotional state because little is known about their temporal characteristics (Hollenstein & Lanteigne, 2014). By applying Kelava et al.'s (2015) synchrony quantification method and simultaneously monitoring the subjective experience of disgust via a rating dial, we were able to compare the real-time changes in physiological synchrony and subjective ratings. Our results revealed that physiological synchrony for neutral and disgusting pictures began to increase in a similar fashion while the participants viewed the black screen that preceded a new picture. We would like to point out that the participants did not know whether the next picture would be neutral or disgusting. The early increase in synchrony stands in line with current empirical results by Bulteel et al. (2014), who reported that the most consistent changes in physiological synchrony can be found at the moment at which the cue, which indicates the upcoming valence of the next picture, was presented. The rating dial, in the present study, was turned after the disgusting picture appeared, and reached its maximum level seconds after the physiological synchrony response, which almost decreased back to zero at this point. The delayed maximum emotional experience might be due to the motor process of turning the dial or emotion regulation strategies such as suppression or reappraisal influenced the temporal difference between physiological synchrony and emotional experience. For example, Gross and Levenson (1993) showed that emotion regulation in disgust has an influence on the physiological but not the subjective level. However, the rating dial values started to increase right after the picture change, and thus, the change itself appeared at the same time when physiological synchrony was at its maximum level. Further, theoretical approaches can also provide potential explanations for our finding. On the one hand, taking the basic emotion approaches or the latent variable model as the underlying theoretical model, the causal chain of emotion elicitation expects the disgusting picture to activate a certain brain circuit, which then causes changes in physiological parameters and subjective experience (Coan, 2010; Ekman, 1992; Tomkins, 1962). During the presentation of the black screen, no obvious stimulus was presented to activate the brain circuit; therefore, the basic emotion approaches and the latent variable model do not seem to be adequate for explaining the early increase in physiological synchrony and delayed emotional experience. On the other hand, the key assumptions of the emergent variable model and the psychological construction approaches might be better able to explain our findings. Throughout the experiment, our participants had already been watching a variable number of pictures before they saw the ones that we included in our main analysis. Thus, one can assume that besides the individual's prior life experiences, a certain preparedness was established throughout the experiment. Due to the context (being in an experimental setting, the possibility that upcoming pictures will be disgusting), physiological parameters begin to change more or less synchronously while participants viewed the black screen, possibly as a form of bodily preparation. Because the subsequent picture is unknown, the process is similar for neutral and disgusting pictures. If the disgusting picture appears, physiological parameters almost automatically change synchronously, appraisal processes start to evaluate the incoming bodily information, the content of the disgusting picture (e.g., a half ripped-off finger), including the affective core (valence and/or arousal), life experience (e.g., working as a physician), and so forth, lead to a delayed conscious awareness of a disgusting state, which leads a participant to turn (or not to turn) the rating dial to a particular level. Thereby, contextual information (picture content) and life experience could influence the emotion-emerging process. For example, a physician might have a different affective core response to the half ripped-off finger than a psychology student, and this difference, amongst others, could explain the interindividual differences. However, the physician him- or herself might also have a different affective core response to a picture of a contaminated toilet compared with a picture of a half ripped-off finger, which, amongst others, could explain the intraindividual differences. If the neutral picture appears, physiological parameters are briefly less synchronously, and appraisal processes begin to be applied to evaluate the incoming bodily information, the content of the neutral picture (e.g., a lamp), including the affective core (valence and/or arousal), and so forth. The participant might have the subjective experience that no emotional state has evolved and therefore, they will not turn the rating dial.

Of course, other explanations are also possible. In both studies (i.e., Bulteel et al., 2014, and the present paper), the changes that occurred prior to the picture could be attributed to some form of orienting response or, in our case, anticipatory anxiety, respectively. In this case, the very brief physiological synchrony around the picture change would denote only a response to the perceptual change and the fulfilled expectation of seeing something disgusting. Hence, no actual experience of disgust would be necessary to evoke changes in physiological synchrony. However, given that the majority of the participants turned the rating dial and reported some sort of disgusting feeling in the situation as well as after the situation, we will assume that the observed effect in synchrony resulted from disgust rather than anticipatory anxiety.[11] In sum, in our view, the early increase in physiological synchrony and the delayed conscious awareness of the subjective experience of disgust support the psychological construction approaches and the emergent variable model.

Methodological issues

Experimental design

In the present study, we used pictures to elicit emotions. Physiological patterns comparable to our study in terms of HR deceleration, increased HRV, and increased SCR have been found in numerous studies of disgust, independent of which emotion-induction method was applied (see Kreibig, 2010, for an overview). The

[11] At this point, we would like to briefly refer to the results of another study of ours. The experimental setup was comparable to the one used in the present study ($N_{Fear} = 34$), but neutral and fear-evoking pictures were shown for 15 s with a black screen of 4 s between the pictures and we had two measurement points (with a time interval of 6 weeks between them). At both measurement points, we were able to reproduce the course of physiological synchrony with regard to the neutral pictures (see Figure A2 in the Appendix). Thus, the short-term decrease does not seem to be some kind of artifact but a pattern reflecting what might be a brief relief. Further, we again found the early increase in physiological synchrony while participants viewed the black screen. The temporal duration of the physiological synchrony response was also comparable to the one found in the present study.

subjective data revealed the largest emotional increase for disgust as compared with fear, anger, or sadness. Even if some of the IAPS pictures that we administered could have evoked emotional states such as fear or anger as well, our manipulation check still supported the successful induction of disgust via pictures. Nevertheless, future experiments should include positive pictures as well to better explain the anticipatory increase in synchrony. For example, future studies could investigate whether the anticipatory increase still appears when participants do not expect the appearance of a negative picture. In one of our prestudies, we showed only neutral and positive pictures and observed the same anticipatory increase in synchrony as we did for disgusting or fear pictures.

Choice of indicator variables for the latent synchrony

As indicators of the latent state synchrony, we chose five pairwise coherence indicators on the basis of their average coherence within delineated regions of interest in the time-frequency plane. Pertaining to disgust, future research is needed to decide, which time-frequency regions have the highest validity and whether other physiological parameters such as saliva or cortisol levels should be included as indicators of the synchrony measure. We would like to emphasize, however, that the applied synchrony measure accounts for many properties of physiological signals, such as non-stationarity, differences in signal power, shape, and so forth. By applying the concept of coherence in delineated regions, we were able to show complex interactions. Note also that these pairwise indicators of coherence have been applied successfully to provide insights into the influence of the cardiopulmonary system on the dynamics of the eye's aberrations (Muma et al., 2010).

Limitations and future directions

The high temporal resolution, the intraindividual quantification approach, and the synchronously acquired online subjective rating are novel and interesting aspects of the current study. Needless to say, our study also has some limitations. First, we focused on physiological synchrony and continuous subjective disgust ratings and did not include the behavioral response system (e.g., facial expression) or brain activity (e.g., electroencephalography). In the future, it would be productive to include more indicator variables in the state-space models and, thus, to provide broader conclusions about the synchrony response in disgust. Second, like numerous studies before, we induced disgust in a controlled laboratory setting via pictures. The artificial emotion-induction method cannot be applied to address the question of whether disgust in the laboratory evokes the same physiological and subjective responses as in a naturalistic setting (Wilhelm & Grossman, 2010). However, body movements and outside stimuli need to be reduced in order to validate the physiological recordings. The use of other physiological sensors that can be more seamlessly connected to the participants (e.g., wearable heart rate sensing devices that use photoplethysmography; Schäck et al., 2015) might make it possible to induce emotions in a more realistic setting in future studies. Third, the temporal characteristics concerning the beginning, duration, and end of the physiological synchrony response to disgusting and neutral pictures certainly depended on our study design. However, they provided a first important step in exploring the temporal resolution of physiological synchrony in general. Future research is necessary to confirm these temporal characteristics for different emotion-induction methods as well as for different emotions. Fourth, in our study, we observed large interindividual differences, a finding that is in accordance with previous empirical results (Bulteel et al., 2014; Hsieh et al., 2011; Mauss et al., 2005). Possible reasons for the interindividual differences are dispositional factors (e.g., disgust sensitivity) and situational factors (e.g., different disgust categories); however, future studies are necessary to investigate the reasons for such interindividual differences.

Final conclusion

In sum, we found a significant difference in physiological synchrony during the onset of a disgusting compared with a neutral picture. Further, we showed that physiological synchrony increased before the actual emotional or neutral picture was presented while the participants watched a black screen that preceded the picture. Our findings support the propositions of Coan's (2010) emergent variable model and of the psychological construction approaches (Barrett, 2013). Physiological synchronization was on average faster and shorter than subjective ratings of disgust. In line with current empirical research results, large interindividual differences in the amplitude, slope, and duration of physiological synchrony were observed. The results of our study might serve as a first step toward a new and emerging research area and provide important fundamental information for future research on physiological synchrony in emotional states.

6. Manuscript C: Person- and Emotion Classification of Fear Using Peripheral Physiological Data

Abstract

The classification of individuals and emotions by use of measured physiological parameters is often restricted to a small sample size, or only a single measurement occasion. The present study is an attempt at taking into account some demanding requirements that typically occur in real world applications: First, individuals were classified from 36 individuals. Second, we considered the challenging case that test and training data were recorded on two measurement occasions separated by an interval of six weeks. Third, the focus was on the person-dependent distinction between fear and a neutral state using nonlinear classifiers. Fourth, in order to estimate the variation of the intensity of an emotion over time, we adapted an extension of support vector machines proposed by Wu, Lin, and Wenig (2004) to estimate time-continuous probability values of the fear state. These values were validated by a continuous subjective online rating measured with a rating dial. Fear was elicited using pictures and film clips. Features derived from the electrocardiogram, the respiration signal, and measured electrodermal activity. Results showed that the classification accuracy of individuals varied depending on the number of measurement occasions and sample size. Film clips produced better classification results compared to pictures for emotion classification. Furthermore, the mean estimated probability values and the subjective ratings of fear showed similar patterns in film clips over time. This unique finding suggests that the probability estimates may serve as an indicator of emotional intensity.

6.1. Introduction

The classification of emotions has been a dynamic and growing research area in psychology over the last decade. The communication efficiency, safety, and convenience of technical equipment such as, e.g., automobiles (e.g., Cai & Lin, 2011) or online learning methods (e.g., D'Mello et al., 2008), would greatly benefit from the capability of sensing and responding to emotional states of its user (Agrawal, Liu, & Sarkar, 2008; Jang et al., 2012). For instance, an online learning program can pause or simplify the questions, if the computer senses frustration in the learner and hence, increase the learner's motivation (D'Mello et al., 2008). Further, a real time classification of the emotional state can be used as a biofeedback method in interventions and therewith help to train the patient's emotional recognition, perception, and clarity in disorders like, e.g. autism (Liu et al., 2008).

The applied modalities to categorize emotional states vary across studies. Some studies utilize facial expression (Valstar & Pantic, 2012; Walecki et al., 2017), gesture (Berthouze et al., 2003; Kleinsmith, Bianchi-Berthouze, & Steed, 2011), voice (Attabi & Dumouchel, 2013; Lee & Narayanan, 2005), or physiological signals (Kragel & LaBar, 2013; Wiem & Lachiri, 2017), while others use combinations of different modalities (Bailenson et al., 2008; Wöllmer, Kaiser, Eyben, Schuller, & Rigoll, 2013). In the present study, we classified individuals and emotions using peripheral physiological data. First, physiological data are less prone to the environment, e.g., illumination conditions and sounds (Jang et al., 2012; Jerritta, Murugappan, Nagarajan, & Wan, 2011), second, they can continuously observe the emotional state in real time (Agrawal et al., 2008; Jang et al., 2012; Jerritta et al., 2011), third, they are independent of overt emotion expression (Agrawal et al., 2008) and hence, can be used to classify emotions in persons that can not speak or do not display facial expressions, fourth, physiological data are directly influenced by the autonomic nervous system and tend to be less consciously mediated by cognitive or social factors compared to facial expressions or speech (Gu et al., 2012; Picard et al., 2001). According to Rani, Sarkar, and Adams (2007), they represent an implicit emotion indicator and might be less interpretable by other humans than by computers. We next highlight some important challenges that appear in emotion classification using physiological signals.

First, in many practical applications, a requirement for any emotion classification method is to perform *person-independent* classification (Jerritta et al., 2011).

Person-independence refers to the source of training and testing data for the classifier, here, training and testing data do not derive from the same person. For instance, person-independent systems become necessary in situations where many different people use the same interface. This is for example the case, when sensing emotions at gaming machines, or in rental cars. With person-independent classification methods, the emotional state of a person can be classified, even though the device did not previously capture training data from this specific individual. However, person-independence is a challenging requirement, since physiological signals strongly vary across individuals, even in the same situational context (Lacey & Lacey, 1958; Kukolja et al., 2014; Rani et al., 2007). This variation impedes the classification accuracy (Kolodyazhniy et al., 2011). Hence, for the implementation of a person-independent classification, a large physiological database is necessary to account for the interindividual differences (Chueh et al., 2012; Jerritta et al., 2011). In a *person-dependent* classification, training and testing data stem from the same person. Person-dependent systems can be applied in private scenarios where the number of individuals that use a specific device is limited, such as, for example, in private computer games, fostering, tutorial systems, or car driving. For instance, a tutorial system may be assumed to have training data from each individual and classifies the emotional state with the help of the individual training data. Person-dependent classification by far outperforms person-independent classification (Kolodyazhniy et al., 2011; Maaoui & Pruski, 2008). Some researchers even question the usability of person-independent classification systems (Chueh et al., 2012). A possible hybrid approach to deal with person-dependent or independent systems can be a combination of person-identification and, in a second step, emotion classification (Kim & Andre, 2008; Li et al., 2016). This is important in practical applications, where it cannot be assumed that the device knows the person for which it performs emotion classification.

A second issue that is highly relevant for emotion classification using physiological signals are the high variabilities or non-stationarities of physiological signals over multiple sessions/recordings. Non-stationarity denotes that the physiological signals statistical properties (e.g. the mean, variance, or covariance function) change over time which propagates on to the extracted physiological features (AlZoubi, Fossati, D'Mello, & Calvo, 2013). As a consequence, there is also a high intraindividual variability in physiological data over a longer time period. Classification algorithms are best when training and testing data originate from the same population, therefore, the non-stationarity represents a great challenge to the development of reliable classification methodologies (AlZoubi et al., 2013; Wei & Jia, 2016). Psychological research has often disregarded the daily variations of physiological signals and measured physiological data only on one single occasion (e.g., Bailenson et al., 2008; Christie & Friedman, 2004; Kolodyazhniy et al., 2011; Kreibig, 2010). This leads to an overestimation of the correct classification rate and does not meet the demands of real world applications (Abdat et al., 2011; R. A. Calvo et al., 2009; Picard et al., 2001). A further challenge stems from the fact that physiological signals are not in general linearly separable in the feature space that is used by the classifier (Zong & Chetouani, 2009). However, studies often concentrate on linear classification methods (Kreibig, Wilhelm, Roth, & Gross, 2007; Rainville, Bechara, Naqvi, & Damasio, 2006; Sinha & Parsons, 1996; Stephens et al., 2010). When a linear separability is not given, nonlinear classifiers outperform linear classification methods and are therefore considered to be a more appropriate method for classification when dealing with physiological signals (Bailenson et al., 2008; Gu, Tan, Wong, Ho, & Qu, 2009; Jang et al., 2012; Kolodyazhniy et al., 2011).

The last aspect which we would like to highlight here is that of accessing the emotional intensity. Some attempts have been made to classify emotion intensity with the help of retrospective subjective data (Agrawal et al., 2008; Rani et al., 2006). However, emotion intensity can strongly vary even during a single film clip (Bailenson et al., 2008), which is not captured through retrospective ratings after the emotion induction. Hence, an online rating method can be a helpful tool for the estimation of emotion intensity. Although, emotional states have successfully been classified in terms of on/off, a more distinguished classification has not been reported, yet. Summing up, physiological signals bear a high variation between individuals. In addition, they have a strong time-dependence which leads to intraindividual variations, e.g. between signals measured at different occasions. Further, the features may not be linearly separable and emotional intensity may vary during a single measurement. As a consequence, it is difficult to build a reliable classification method that fulfills the demands of real world applications. Since these factors have often been ignored correct classification rates were overestimated (Picard et al., 2001).

In the present paper, our contributions are threefold: First, we considered the classification of individuals and emotions using peripheral physiological signals. To the best of our knowledge, the inclusion of person classification has not been reported in psychological research so far. It gives an indication for the intraindividual stability of physiological data over time, which has been a major research topic in psychophysiology (Hinz, Hueber, Schreinicke, & Seibt, 2002; Rani et al., 2007). Therefore, we collected data of 36 participants on two measurement occasions separated by a time interval of six weeks. In this way, we included intraindividual variations of physiological data over time, which yielded a more realistic picture in terms of classification accuracy. Second, we concentrated on the person-dependent distinction between fear and neutral states using nonlinear classifiers. Fear was induced with pictures and film clips. Existing research has often restricted the setup to only one measurement occasion (e.g., Kreibig et al., 2007; Rainville et al., 2006; Stephens et al., 2010). Only a few have collected data in a repeated measurement design (e.g., Agrawal et al., 2008; R. A. Calvo et al., 2009; Kukolja et al., 2014; Picard et al., 2001; Rani et al., 2006). As a drawback, these studies have only used 1 to 15 participants. In the present study, data derived from two measurement occasions with 36 participants. Third, in order to estimate the variation of the intensity of an emotion over time, we adapted an extension of T. F. Wu et al. (2004) for one of the nonlinear classification methods applied in our study. The extension provided an estimate for the probability of an observation to belong to a certain class. This allowed for a more continuous distinction between the emotional states. The computed probabilities were compared to a subjective online rating that was measured with a continuous rating dial. To the best of our knowledge, this approach has not been applied in emotion classification research, so far.

The paper is organized as follows: First, we discuss some important theoretical models of emotion that have been applied in emotion classification research. Second, we give a short overview of the research on this area including classification methods and related empirical work. Following, we point out our research aims, describe in detail the proposed methodology, and provide and discuss the obtained results both for person- and for emotion classification.

Theoretical background

The underlying theoretical model of emotions to be used in emotion classification research is a topic of ongoing controversial debate. On the one hand, the basic emotion approach assumes unique and distinctive response patterns of physiological, behavioral, and subjective channels for different emotions like fear, anger, joy, or disgust (Ekman, 1992; Izard, 1977; Levenson, 1994; Tomkins, 1962). This approach has been applied in various studies (Kragel & LaBar, 2013; Picard et al., 2001; Stephens et al., 2010; Walecki et al., 2017). On the other hand, the dimensional approach arranges different emotions in a two dimensional space spanned by valence (positive/negative) and arousal (high/low; Barrett, 2006a; Barrett & Russell, 1998; Russell, 1980, 2003) and has been applied for emotion classification as well (Gu et al., 2012; Haag, Goronzy, Schaich, & Williams, 2004; Jatupaiboon et al., 2015; Swangnetr & Kaber, 2013; Walter et al., 2013). The results supporting either model are mixed: Some support the idea of emotion specific subject independent response patterns (Ekman et al., 1983; Friedman & Kreibig, 2010; Levenson, 1992; Rainville et al., 2006; Stephens et al., 2010), others question response specificity in emotion (Cacioppo et al., 2000; Ortony & Turner, 1990; Russell, 2009). Further, the exact number of basic emotions and the distinction between basic-emotion and non-basic emotion are not conclusively determined and are thus the subject of an ongoing scientific debate (Barrett, 2006a; Grandjean et al., 2008; Ortony & Turner, 1990; Russell, 2009). For instance, Ekman and Cordaro (2011) proposed a set of seven basic emotions: anger, fear, surprise, sadness, disgust, contempt, and happiness, while Izard (2011) recommended the use of six basic emotions (interest, enjoyment/happiness/contentment, sadness, anger, disgust, and fear). Pertaining to the dimensional approach, some researchers consider the two-dimensional space to be an insufficient model for emotion classification (Bradley & Lang, 1994; Fontaine et al., 2007). In addition, there are differing scientific opinions on the interpretation of the axes themselves: Some theorists claim an inverse relationship between positive and negative emotions, while others assume them to be independent (Mauss & Robinson, 2009). Further, the reduction of an emotional state to two or three dimensions is also criticized by some researchers (DeSteno, Petty, Wegener, & Rucker, 2000; Lerner, Dahl, Hariri, & Taylor, 2007; Rosenberg & Ekman, 1994). Kragel and LaBar (2013) compared the discrete and dimensional approach using multivariate physiological parameters. Their results supported the discrete approach (see also Chanel, Kierkels, Soleymani, & Pun, 2009; Lench et al., 2011; Stephens et al., 2010). Although we are fully aware of the shortcomings, in this work, based on the above results, we treat fear as a distinctive emotional state that can be divided in different intensity levels.

Classification methods

All studies reported in this short overview used physiological data as underlying modality. One of the most commonly applied linear classifiers are the linear discriminant functions (e.g., Christie & Friedman, 2004; Rainville et al., 2006; Sinha & Parsons, 1996), in which a hyperplane divides the input space into two classes. Extensions for multiple classes have been developed and the parameters of the discriminant functions can be estimated using least squares, among other approaches (Bishop, 2006). Several studies demonstrated that non-linear classifiers like k-nearest neighbors (KNN; Cover & Hart, 1967), support vector machines (SVM; pioneered by Vapnik, 1998), or neural networks outperform linear classification methods when the features are not linearly separable (Bailenson et al., 2008; Gu et al., 2009; Jang et al., 2012; Kolodyazhniy et al., 2011). The KNN is a non-parametric method for classification and was applied in a variety of studies (e.g., Kortelainen et al., 2012; Li et al., 2016; Zhou, Qu, Jiao, & Helander, 2014). A new data point is classified according to the labels of its k-closest training points. Using Bayes theorem, new data is assigned to the class with the highest posterior probability (Bishop, 2006). The SVM determines a hyperplane with a maximal margin to the nearby data points (called support vectors) to separate the classes and has recently become very popular in classification research (e.g., Abdat et al., 2011; Cheng, Chen, & Wang, 2012; Kragel & LaBar, 2013; Wei & Jia, 2016). To overcome the limitation of linear separability, different kernels like the Gaussian kernel can be applied (Bishop, 2006). A third framework of non-linear classifiers is based on neural networks. Amongst these, the Multilayer Perceptron (MLP, Haykin, 1999) is one of the most popular neural network methods. MLP is a feedforward neural network consisting of multiple layers. These layers comprise an input layer, one or more hidden layers, and an output layer. The connections among the neurons are weighted according to their correlation. Classification can be achieved by learning weights of the network using training data (Bishop, 2006). The MLP approach has been applied in different studies (e.g., R. A. Calvo et al., 2009; Kolodyazhniy et al., 2011; Wagner, Jonghwa, & Andre, 2005). It is important to highlight that amongst these methods, there is no single best classification algorithm but rather the characteristics of the data determine which method should be used (Kim & Andre, 2008). Results are generally reported in form of a correct classification rate (CCR) which gives the overall percentage of correctly classified elements. In the present study we apply the KNN and SVM classifiers.

Related empirical work

Classification of individuals

The classification of individuals via Electrocardiographic (ECG) signals has been an active and growing area of research in recent years (e.g., Chan, Hamdy, Badre, & Badee, 2006; da Silva, Fred, Lourenco, & Jain, 2013; Lourenco, Silva, & Fred, 2012). Practical applications include biometric services for person verification. In terms of methodology, neural networks (CCR = 86.10%; Boumbarov, Velchev, & Sokolov, 2009), SVMs (CCR = 97.54%; Ye, Coimbra, & Kumar, B. V. K. V., 2010), and the KNN (CCR = 97%; Venkatesh & Srinivasan, 2010) were applied. It is necessary to mention here that the majority of the studies classify data recorded on one measurement occasion, i.e., training and test data come from the same population (e.g., Boumbarov et al., 2009; Fratini et al., 2013; Ye et al., 2010). Multiple occasions are however necessary to account for the intraindividual variabilities over time (Chan et al., 2006; da Silva et al., 2013). For instance, da Silva et al. (2013) highlighted the variability by comparing person classification within one and between two measurement occasions. For several subjects, the heartbeat waveforms changed significantly between the two occasions, even though both measurements were taken in the same non-emotional setting. The recognition rate was higher if only data recorded on a single day was used (for similar results see also Odinaka et al., 2012). The results reflect the daily variability of physiological data and emphasize the necessity of including multiple measurement occasions when performing classification. Further, the data recording in the former studies usually took place in a calm environment without any expectation of an emotional event or excitement on the part of the participants. Any kind of physical or psychological activation or excitement can decrease the correct recognition rate of individuals (Matta, Lau, Agrafioti, & Hatzinakos, 2011; Odinaka et al., 2012). Hence, it is questionable, if similar classification results can be achieved, if the individual expects emotion evoking stimuli. Studies that consider psychological stress factors like emotions in their person classification are rare (Odinaka et al., 2012), more research needs to be conducted in this field. Summing up, most studies use data that is recorded in a calm and unemotional session where training and test data come from a single measurement occasion. Studies that include psychological stressors and multiple measurement occasions are rare but necessary to realistically address the challenges imposed by real world applications.

Classification of fear

Fear plays an important role in various human-machine interactions, for example in task performance (Rani et al., 2007). Kreibig (2010) concluded in her review that, combining all studies, fear is associated with cardiac acceleration, increased electrodermal-, and respiratory activity. Further, a decreased heart rate variability (HRV), increased P-wave amplitude, and decreased T-wave amplitude were reported in most studies. A decrease in heart rate (HR) was also found by studies presenting pictures (e.g., Bernat, Patrick, Benning, & Tellegen, 2006; Codispoti & de Cesarei, 2007) or film clips (e.g., Fredrickson & Levenson, 1998). When using more realistic emotion induction methods such as sudden announcement of an exercise or outage of the light, a decreased electrodermal activity could be observed (e.g., Krumhansl, 1997; Stemmler, 1989, for an overview see Kreibig, 2010).

Numerous studies of emotion classification have induced fear on a single measurement occasion (e.g., Christie & Friedman, 2004; Kolodyazhniy et al., 2011; Kreibig et al., 2007; Lisetti & Nasoz, 2004; Maaoui & Pruski, 2008; Rainville et al., 2006) and with a small sample size from 4 to 12 individuals (e.g., C.-Y. Chang et al., 2009; Gouizi et al., 2011; Jang et al., 2012). The rate of correctly classifying fear varied from study to study, using a linear classifier the CCR reached 57.4% (Jang et al., 2012), for the KNN, the CCR was 81% (Lisetti & Nasoz, 2004), for neural networks the CCR reached 83% (C.-Y. Chang et al., 2009), and the SVM achieved 91% (Maaoui & Pruski, 2008). Note, however, that a direct comparison is not possible, since all studies used different designs and the number of induced emotions varied. Further, a single measurement occasion possibly leads to an overestimation of the CCR, since they do not include intraindividual variations of physiological data (R. A. Calvo et al., 2009; Picard et al., 2001). Hence, our study constitutes an important benchmark for classification accuracy of fear in real world applications. Summing up, though the majority of the studies found sympathetic activation during a fear response, empirical results regarding the physiology of fear are inconsistent. Existing research only classified fear using a single measurement occasion. Studies with a large sample size and two measurement occasions are missing. Hence, the present study provides an important insight in classification including intraindividual physiological variation.

Continuous classification of emotions

The classification methods that have been applied so far are binary in terms of their decision, i.e., they categorize whether an emotional state is present or not (zero/one). Attempts have been made to classify different intensity levels of an emotional state. Recently, Walecki et al. (2017) applied successfully variable-state Latent Conditional Random Fields to detect intensity levels of facial emotion expressions. Rani et al. (2006) induced anxiety, engagement, boredom, frustration, and anger with computer based cognitive tasks. They divided the self-reported output states in three different intensity levels (low, medium, and high), classification was then conducted on each affective state individually with a person-dependent classifier. Amongst other classifiers, they compared the KNN with the SVM (non-linear kernel). Classification accuracy varied between 70% and 89%, depending on the emotion and classifier. The SVM outperformed the KNN, for instance, the recognition rate for anxiety was 80% for the KNN and it reached 89% for the SVM (see also Agrawal et al., 2008). Summing up, retrospective, self-reported data or facial expressions were used to classify different intensity levels of emotion. To our knowledge, there is no study which used continuous, online self-reported data or a classification method that reports different probabilities on the certainty of the decision for an emotion using physiological input data.

The present study

The focus of our research paper lay on three major aspects: First, we concentrated on the classification of individuals. Here, we wanted to investigate the influence of the number of measurement occasions as well as that of a larger sample size (more individuals) on the classification accuracy. Second, we focused on the distinction between fear and and a neutral emotional state. Therefore, we compared the classification results obtained from the KNN with those of the SVM as well as pictures with film clips in terms of the induction methods. Through our study design, with a larger sample size and two measurement occasions, we considered realistic conditions that better meet the requirements of real world applications, compared to previous research. Third, we continuously estimated different probabilities for the fear state over time and compared the results with a continuous self-reported rating of emotion intensity. By comparing the probability estimates with continuous self-reported data, we wanted to analyze, if a higher probability estimation for the fear class is accompanied by a higher subjective fear rating and vice versa. Prior studies only used retrospective subjective experience data for emotion intensity. Moreover, the research of classifying probabilities for a certain emotional state has, to our knowledge, not been conducted so far.

6.2. Method

Participants

In total, 42 participants took part in our study. From the beginning, a sample of six participants had to be excluded from the analysis because they had a diagnosis of mental illness, they smoked before the experiment, or a technical breakdown occurred (e.g., computer froze). Therefore, 36 participants (20 female, 16 male) were included in the analysis (M = 23.66, SD = 4.09, range = 19-38 years). The majority of the participants were students who could collect course credit for their participation. All participants filled out consent forms. The experiment was approved by the local ethic committee.

Emotion Elicitation

Various emotion induction methods have been applied in emotion classification research: pictures (e.g., Gu et al., 2012; Haag et al., 2004), film clips (e.g., Bailenson et al., 2008; Leng, Lin, & Zanzi, 2007), music (e.g., Kim & Andre, 2008; Wagner et al., 2005), or imagination of experienced emotional situations (e.g., AlZoubi, Calvo, & Stevens, 2009; R. A. Calvo et al., 2009; Picard et al., 2001). In the present study, we chose film clips, because they elicit emotions effectively, and can be standardized across participants (Hewig et al., 2005; Rottenberg, Ray, & Gross, 2007; A. Schaefer, Nils, Sanchez, & Philippot, 2010). Fernandez et al. (2012) showed that fear films are able to produce a significant change in physiological data. Further, we wanted to distinguish different probability levels of the fear class. Watching film clips elicits a varying level of an emotional state (Bailenson et al., 2008), hence this method seemed appropriate for the purpose of this study. We wanted to test our findings against undynamic yet standardized fear inducing stimuli like pictures (Lang et al., 2008, 1993). The emotions that are elicited by pictures generally have a shorter duration and less intensity fluctuation than film clips. The International Affective Picture System (IAPS; Lang et al., 2008) is categorized through the dimensions valence and arousal (Lang et al., 2008, 1993), but attempts have been made to label the pictures according to discrete emotional categories, showing that fear can also be successfully elicited using the IAPS (Barke, Stahl, & Kröner-Herwig, 2012; Mikels et al., 2005). Lench et al. (2011) showed in their study that pictures are the most effective elicitation method for discrete emotions. Hence, we have chosen pictures as a second emotion elicitation method.

Baseline block

There were three different blocks throughout the experiment in both measurement points (MPs; see Figure 6.1). During the *baseline block*, eight neutral IAPS pictures[12] were displayed in randomized order. Pictures for the baseline block were selected based on their relatively neutral rating of valence and low rating of arousal (Lang et al., 2008). Further, one neutral film clip explaining the basics of genetics was shown at MP 1 (8 min 26 s), a film clip about encoding the DNA in humans at MP 2 (7 min 26 s).

Picture block

During the *picture block*, two affective pictures for each category: violence[13] (e.g., a gun pointed at the participant), accident scenes[14] (e.g., car accident), and general threat[15] (e.g., a sinking ship) were displayed. Two neutral pictures[16] (e.g., a book) were added to ensure that participants have a relaxing phase in between. Pictures for each MP were displayed in the following categorical order: general threat, accident, neutral, and violence. The pictures within each category were randomly assigned. The majority of the pictures were taken from IAPS but to counterbalance the different categories some pictures were added. All pictures in the experiment

[12] IAPS picture numbers of the baseline block MP 1:7006, 7009, 7010, 7025, 7035, 7150, 7217, 7235; IAPS picture numbers of the baseline block MP 2: 5390, 5720, 7020, 7211, 7031, 7052, 7175, 7491.

[13] IAPS picture numbers of violence pictures for MP 1: 6260, 6560; IAPS picture numbers of violence pictures for MP 2: 6350, 6570.1.

[14] IAPS picture numbers of accident pictures for MP 1: 9901 and one picture of a train accident; IAPS picture numbers of accident picture for MP 2: 9903 and one transporter accident.

[15] IAPS picture numbers of general threat picture for MP 1 and MP 2: 9050, 9600; IAPS picture numbers of general threat pictures for MP 2: 9630, a picture of people running away from a tsunami wave

[16] IAPS picture numbers of neutral pictures for MP 1: 7002, 7090; IAPS picture numbers of neutral pictures for MP 2: 7040, 7490

had the same quality (72 dpi) and were each displayed for 15 s (see also Haag et al., 2004) with a 4 s black screen between each picture.

Film block

During the *film block*, four film clips (two fear inducing and two with neutral content) were shown for each MP. The fear inducing film clips were chosen from the FilmStim database (A. Schaefer et al., 2010). For MP 1, a scene of *The Shining* (4 min 26 s; also recommended by Rottenberg et al., 2007) and a scene of *The Blairwitch Project* (4 min 5 s) were shown. Additionally, two neutral film clips were displayed in between, to give the participants a short period to calm down: a scene of a documentary explaining the making and roasting of coffee (3 min 30 s) and a scene of a documentary about a potato field on a typical German farm (4 min 2 s). For MP 2, a scene of *Misery* (5 min 7 s) and *Scream 1* (6 min 57 s) were shown. The two neutral film clips included a scene of a documentary about an environment-friendly house (3 min 37 s) and a scene of a documentary about producing a newspaper (3 min 52 s). The order of the film clips was fear, neutral, fear, neutral, whereby the fear and neutral film clips where randomly assigned within each MP.

Questionnaires

The *Differentielle Affekt Skala* (DAS; Merten & Krause, 1993) is a German adaptation of the *Differential Emotions Scale* (Izard et al., 1974). The DAS is a list of 30 adjectives and measures ten emotional states *interest, joy, surprise, grief, anger, disgust, contempt, fear, shame,* and *guilt*. Each emotional state is captured by three adjective (e.g., *fear* by: "frightened", "fearful", "scary"). Participants were asked to rate their current feelings ("How do you feel right now?") on a five point Likert scale (1 = *not at all* to 5 = *very strong*) after each block for MP 1 and MP 2.

Physiological recording equipment

ECG, respiration, electrodermal activity (EDA), and rating dial were measured with Biopac MP 150 System and AcqKnowledge 4.2 Software (Biopac Systems Inc, 2011). The data was sampled at a frequency of 250 HZ.

Electrocardiogram

The ECG was recorded using pregelled silver/silver-chloride electrodes attached to the body in a modified chest LEAD III configuration. The RA lead was placed at the beginning and the LA lead at the end of the sternum. The LL lead was attached to the left margin of the chest, on the third rib bottom up. Electrodes were connected to the ECG100C amplifier of the MP150 polygraph. To remove measurement artifacts that were induced when subjects moved during the ECG recording, we applied an artifact filter by Strasser et al. (2012). The filter operates in the stationary wavelet domain. In this domain, an ECG signal is decomposed into a set of so-called wavelet coefficients, which represent information on different time-scales. Using this decomposition, the artifacts can be estimated and subsequently subtracted from an ECG signal. By applying the algorithm an artifact cleaned ECG signal was obtained.

Respiration

Respiratory activity was recorded with the effort transducer TSD201, a stretch sensitive belt placed below the sternum. The participant was supposed to exhale completely, then the belt was tightened when the chest was at its smallest diameter. The respiration transducer was attached to the subject in a sitting position, which was maintained. Next, the belt was connected to the RSP100C amplifier. No preprocessing was done except for applying a bandpass filter with cutoff frequencies of 0.05 Hz and 10 Hz, respectively.

Electrodermal activity

EDA was recorded using 8 mm diameter silver/silver-chloride-electrodes. The electrodes were filled with isotonic gel (0.5% saline in a neutral base) to provide a continuous connection between the electrodes and the skin and were placed on the palm of the non-dominant hand. The electrodes were connected to the GSR100c module, amplified with a gain set at 5 μS/V and recorded using a constant voltage of 0.5 Volt. In the preprocessing step, a butterwoth lowpass filter (Oppenheim, Willsky, & Nawab, 1983) with a cutoff frequency at 20 HZ was used to filter out high frequency noise and the inductively coupled power hum.

Rating dial

For each MP and each block (neutral, picture, and film-clip), the participants rated their emotional state with a rating dial. The online rating method does not have an impact on the emotional experience (Mauss et al., 2005). The potentiometer was attached to a UIM100C-modul and had a scale from 0 (*not frightening at all*) to 10 (*very frightening*) and an incoming voltage range from 0 to 5 Volt. The dial was placed in a hemicycle on a 20 × 13 × 8 cm box.

Procedure

Participants were invited twice to the laboratory, separated by a time interval of six weeks. Both MPs had the same course of events, only the stimulus material varied. At the beginning, participants were welcomed in the laboratory and seated in a normal office chair in a 4 × 6 m laboratory room with a distance of 0.5 m to the computer screen. They were told that they would see different pictures and film clips that can evoke fear. The participants were supposed to rate their perception of fear continuously on a rating dial. Thereby, 0 meant *not frightening at all* and 10 meant *very frightening*. The experimenter pointed out that they should only rate their feeling of fear, they should not rate any other negative affect. They could stop the experiment at any time and still get their course credit. Further, they were reminded to avoid unnecessary movements and to stay as calm as possible. After the instructions, the participants were connected to the physiological measurement devices.

The experiment can be divided into three parts (see Figure 6.1). All participants started with the baseline block. The order of pictures or the film-clip at the beginning of the baseline block was randomized for each participant. We always started with the neutral material in order to assure that the baseline measurement is not influenced by previous affective material. Between each block, the participants filled out questionnaires on the computer. After the baseline block, the participants randomly watched first the picture block and then the film block or vice versa for both MPs. In this way, sequence effects of the presented material were ruled out.

Throughout the experiment continuous physiological and subjective data were measured and additionally, retrospective ratings were taken via questionnaires. The experimenter was in the same room as the participant, to monitor the physiological experiment, however they were separated by a room divider. For programming the experiment and the selection of pictures and film clips, we used the software Matlab R2011b (MathWorks, 2011).

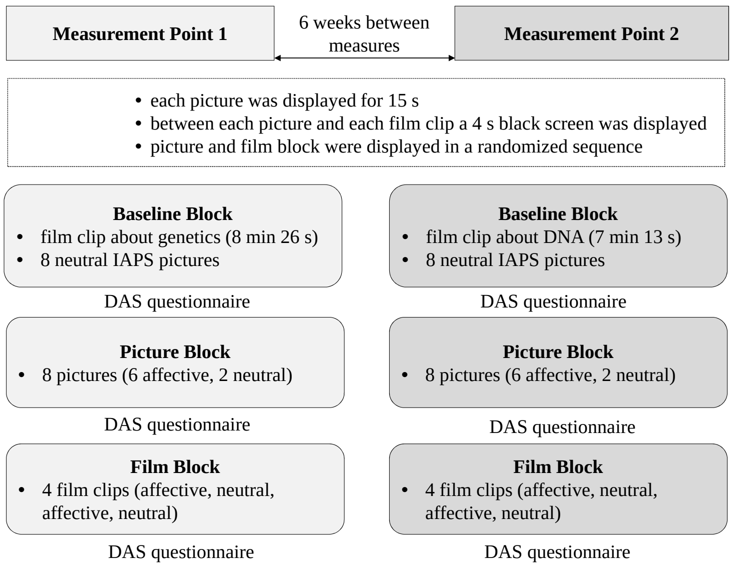


Figure 6.1. The course of the experiment at measurement point 1 and 2. DAS stands for the *Differentielle Affekt Skala* (Merten & Krause, 1993), IAPS for the International Affective Picture System (Lang et al., 2008).

Applied classification algorithms

Classification is a fundamental research topic of high interest with numerous applications, e.g. in the areas of machine learning or medical science. It deals with the problem of determining to which of a set of predefined categories a new observation belongs. There are two general types of classification algorithms. In *supervised* classification, the new data is categorized based on a set of training data whose class membership is known, whereas *unsupervised* classification tries to categorize the data without having any training data available (Dougherty, Kohavi, & Sahami, 1995). In the following, two supervised classifiers are described in more detail which are used for the evaluation of our experiments.

KNN

The KNN classifier is a non-parametric classification method, that was first introduced by Cover and Hart (1967). It is a method of choice when there is little or no prior knowledge about the distribution of the data. The KNN exploits the idea that nearby objects are likely to have similar properties. The training phase consists solely of storing the feature vectors with their class labels. For the classification of a new, unlabeled vector, the KNN-algorithm first chooses its k nearest vectors among the set of training vectors (Fix & Hodges, J. L.,Jr., 1989). A commonly used distance metric is the Euclidean distance, where the distance between two vectors $\mathbf{x}$ and $\mathbf{y}$ is given by

$$d_{\text{euclidian}}(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{d=1}^{D} (x_d - y_d)^2}. \tag{6.1}$$

Here, D denotes the dimension of the feature vector (Fix & Hodges, J. L.,Jr., 1989; Wang, Neskovic, & Cooper, 2007). The labels of these k nearest neighbours are analyzed and the label which appears most frequently is assigned to the new feature vector. A drawback of the KNN is the fact that its computational complexity grows rapidly for huge data sets (Guo, Wang, Bell, Bi, & Greer, 2003). However, its simple implementation and lack of required prior knowledge of the data makes the KNN a popular classification algorithm.

SVM

The SVM algorithm was proposed by Vapnik (1998). Its basic idea is to divide the feature space by a hyperplane in such a way that the feature vectors of each class are optimally separated from each other. There are many approaches to this task. The optimal solution describes a hyperplane which results in the largest separation, called margin, between the classes (Bishop, 2006). This approach, however, supposes that the feature vectors are linearly separable. This restriction can be circumvented by applying the *kernel trick* (Bishop, 2006): If the data cannot be linearly separated in the input space, a so-called kernel function is used to map the features to a higher dimensional space where the features are linearly separable. For the evaluation, in this paper, the Gaussian kernel was used. The Gaussian kernel is an exponentially decaying function in the input feature space which has its maximum value at the position of the support vector and which is used to calculate the distance of a new feature vector to one of the support vectors. The smaller the distance between the vectors becomes, the higher becomes the output of the kernel function and the importance of this support vector for the classification of this feature increases. This procedure is done for every support vector. The SVM uses a weighted linear combination of the resulting kernel functions for the classification while the weighting is determined by the importance of the respective support vector for the classification. The central idea is that the superposition of kernel functions representing each of the two classes draws the boundary between them. The smaller the distance of a feature to a certain kernel, the higher is the probability that they share the same class label. The Gaussian kernel is able to adjust to the data very well, that is why it can be used to recognize even very complex patterns. However, the price to be paid is a higher computational complexity (Bishop, 2006).

Since the SVM predicts only class labels without any membership probability, we adapted a method that was introduced by T. F. Wu et al. (2004) which provides an estimate for the probability of an observation to belong to a certain class. This estimate is obtained by first comparing the input data with the pairwise class boundaries which are calculated based on training samples. These estimates of the pairwise comparisons are combined and the results are iteratively optimized until a probability estimate for all classes is obtained. This extension is implemented in the Matlab LibSVM (C.-C. Chang & Lin, 2011) toolbox, which is applied in our experiment.

Data analysis

Feature extraction

Emotion classification studies have often reduced the extracted features to a subset which is most effective for analysis. Thereby, the number of features used for the analysis varied from study to study. In our paper, we a priori selected nine features, which are sensitive to capture emotional activity (Gramann & Schandry, 2009). Further, physiological signals are corrupted by external influences such as measurement noise and artifacts. Especially artifacts caused by body movements are critical. Therefore, each signal passed a preprocessing step to reduce the noise and motion artifacts, as described above. Following, the features were extracted from the preprocessed signal. After the extraction, the feature signals were split into segments of 15 s duration. For each segment, one feature vector was calculated.

Cardiovascular features. The cardiovascular features were extracted from the ECG. The analysis of the ECG signals is structured as follows: First, the R-peaks of the ECG waveform were detected with the method of Chen, Chen, and Chan (2006). Second, from these detected R-peaks the *Interbeat Interval* (IBI) was calculated by analyzing the time difference between two successive R-peaks. The related HR can be calculated by using the formula HR = 60/IBI. In the present study, we only used the IBI as feature, since the IBI and the HR are linearly connected with each other and there was not a gain of information by using both features. The HRV was calculated from the variance of the IBIs. Applying a constrained minima and maxima search, the P, Q, S and T- points were detected and the *QT-Time* and the *QT-Amplitude* were calculated for emotion classification.

For the person identification task, we classified the individuals based on the morphology of the ECG wave. Therefore, we calculated a number of features which describe the amplitudes and timing intervals of the ECG wave. As timing intervals we used the *QT-Time*, the *QS-Time*, the *PQ-Time*, and the *ST-Time* which were normalized to the current HR. We used relative feature amplitudes so as to be independent of the absolute values which vary, depending on the measurement setup. In particular, the *QT-Amplitude / PQ-Amplitude*, the *QT-Amplitude / QS-Amplitude* and the *QS-Amplitude / PQ-Amplitude* were used for the person classification task. Similar procedures for person identification were applied in Biel, Pettersson, Philipson, and Wide (2001) or Shen, Tompkins, and Hu (2002).

Respiratory features. For feature extraction, a peak detection algorithm searched for the inhalation peaks and the exhalation minima. The *Respiration Cylcle Time* (RCT) was then calculated by the differences between two inhalation peaks. The *Respiration Frequency* (RF) can be calculated by the formula RF = 1/RCT. Because the RCT and the RF are strongly correlated only the RCT was used as feature. The *Inspiratory Depth* was calculated by the distance between the inhalation peak and the exhalation minimum of the respiration signal.

Electrodermal features. The features which were calculated for each segment were the EDA *Slope*, the *Time of Positive Increase* of the EDA, and the *Time of Negative Increase* of the EDA. We only used the derivative signals of the EDA because these are less influenced by previous segments than the EDA level itself since the trend was canceled out.

Statistical analysis

In order to perform emotion and subject classification with the methods described above, feature vectors must be formed. From the cardiovascular, respiratory, and electrodermal features we constructed a feature vector of nine features for emotion classification. From the cardiovascular features we constructed a feature vector of seven features for classification of individuals. The features selected for emotion identification were assumed to be sensitive to the presented emotional content and the features for classification of individuals were person dependent. There were different measures with different value ranges, and different units in the vectors. To weight the influence of each feature equally, we normalized the features by their robust estimates of the mean and variance with the following formula:

$$F_{norm} = \frac{(F - median(F))}{1.4826 \cdot mad(F)}, \tag{6.2}$$

where F is the feature, the median is a robust estimate of the mean, the median absolute deviation (mad) is a robust estimate of the standard deviation, and the factor 1.4826 provides consistency when assuming a Gaussian distribution of F (Zoubir et al., 2012). Robust estimates were used, since they are insensitive towards outliers in the feature vectors. The classification was now performed with both classifiers, i.e., the KNN and the SVM with a Gaussian Kernel.

Classification of individuals

For the classification of individuals we only used baseline film clips, because the person had a neutral stimulus input over a relatively long period and hence, more data points were available. Two scenarios were constructed: In the first scenario, the individual was classified based on the data of one MP. The classification accuracy was evaluated with a two-fold cross validation, whereby the baseline film clip was split in two equal parts, each part was assigned to one fold. Hence, training and testing data derived from the same MP. In the second scenario, both MPs were used for the classification. Again, a two-fold cross validation was used, this time the baseline clip of each MP was assigned to different folds.

Classification of fear based on pictures

The pictures were shown for 15 s in the experiment. Therefore, one picture represents one feature vector in our evaluation. The first two pictures of the baseline were left out of the analysis to reduce the effect of the initial excitement of the participants. We conducted a two fold cross-validation, where the pictures of one MP are assigned to one fold. Hence, the classifier is trained with the pictures of one MP and tested against the other MP.

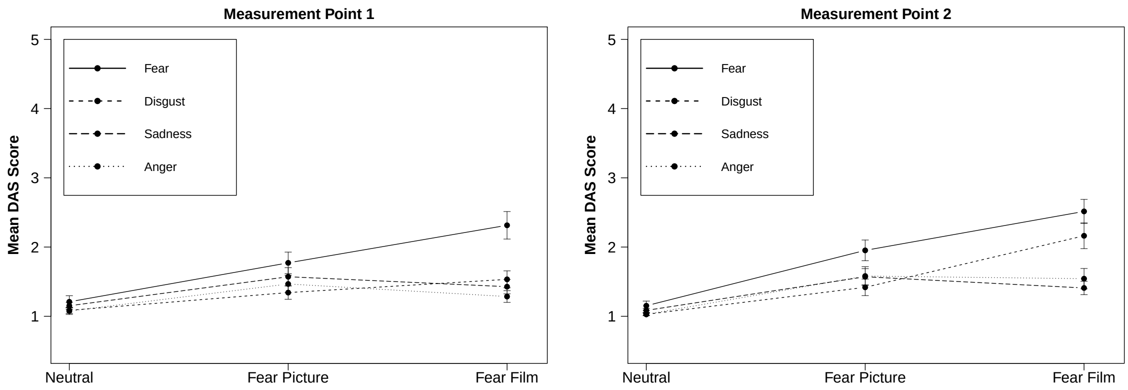
Classification of fear based on film-clips

For the classification of fear based on film clips, we used the probability extensions for the SVM that were described above. The KNN was applied in its usual form. The information of the probability is an indicator for the reliability of the classification decision and facilitates the interpretation of the results over time. As for the pictures, we applied a two fold cross-validation, and the features were calculated for 15 s frames. We trained the classifier for each person separately. We then combined the folds as follows: the middle 4 min of the initial baseline film clip and one of the emotional film clips (*The Shining* or *Blairwitch Project* for MP 1 and *Misery* or *Scream* for MP 2). There were four possible video combinations: 1. Cross-validation with *The Shining* and with *Misery.* 2. Cross-validation with *The Shining* and with *Scream.* 3. Cross-validation with *Blairwitch Project* and with *Misery.* 4. Cross-validation with *Blairwitch Project* and with *Scream.*

6.3. Results

Manipulation check

To ensure the elicitation of fear through our stimulus material, we analyzed the retrospective, subjective data of the DAS. Figure 6.2 depicts the mean DAS scores for fear, disgust, sadness, and anger after the neutral stimuli, the fear pictures, and the fear film clips for MP 1 and MP 2. The neutral stimuli did not elicit any negative emotions (see Figure 6.2). Fear ratings were higher in measurement point 2 (see Figure 6.2(b)), during MP 1 the increase of fear was only slightly higher than for sadness (see Figure 6.2(a)). Film clips produced higher fear rating than pictures. Yet, during MP 2 disgust ratings were almost as high as fear ratings (see



(a) Mean DAS scores on measurement point 1 (b) Mean DAS scores on measurement point 2

Figure 6.2. Averaged retrospective subjective ratings of the *Differential Affect Scale* (DAS; Merten & Krause, 1993) after neutral stimuli, fear pictures, and fear film clips on measurement point 1 and 2. Fear ratings were higher using film clips in both measurement points.

Figure 6.2(b)). Summing up, the emotion induction with film clips seemed to be more successful than using pictures.

Classification of individuals

The random chance of correct classification was 2.78% for each individual. The random chance of correct classification is an important benchmark for the evaluation of a successful classification and depends on the number of participants (the higher the number of participants the lower is the random chance of correct classification). The classification of individuals was conducted within MP 1 and between MP 1 and MP 2. Figure 6.3 depicts the confusion matrices for the classification accuracy of the SVM classifier. The confusion matrix can be interpreted as follows: Each column represents a real person while each row represents a predicted person. The diagonal shows the probability for a correct classification between 0 and 1 for each individual, false alarms are displayed outside the diagonal.

The results largely supported our assumptions. First, Figure 6.3(a) displays the confusion matrix for the classification within MP 1. For some individuals like no. 4, 13, or 19, the CCR was almost 100%, while for other individuals like no. 11 or 23, the CCR was between 10 and 20%. Hence, there is a high interindividual variability in the classification accuracy. In total, 56.70% (54.53%) of all individuals were correctly classified using the SVM (KNN). The results were clearly above the random chance of correct classification.

Second, Figure 6.3(b) depicts the confusion matrix when training data derived from MP 1 and testing data from MP 2 of our experiment. Again, for some individuals like no. 4, 6, 15, or 19, the CRR was close to 100%. However, compared to Figure 6.3(a) there were more individuals with a CCR below 10% (e.g., no. 9, 13, 16, or 18) as well as more false alarms. In total, 26.93% (23.16%) of all individuals were correctly identified using the SVM (KNN). The CCR declined from 56.70% to 26.93%, if data recording derived from two different measurement occasions and hence, two different recording setups. The impact of daily variations in physiological data on classification accuracy is apparent.

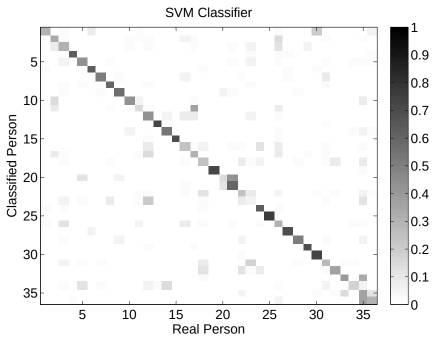
Since our results have lower CCR than reported in other studies, we want to exemplarily show, which influence the sample size can have on the reported CCR. Therefore, we used the baseline data from a similar experiment with the same feature vectors, time frames, and classifiers, but with three measurement occasions and only five participants. Two measurement occasions were used for training- and the third occasion for testing data. Figure 6.3(c) displays the confusion matrix of the person identification using the SVM. Random chance of correct classification was by 20% for each individual. In this task, CCR reached 87.30% for the SVM and 84.8% for the KNN, respectively.

Classification of the fear state in pictures

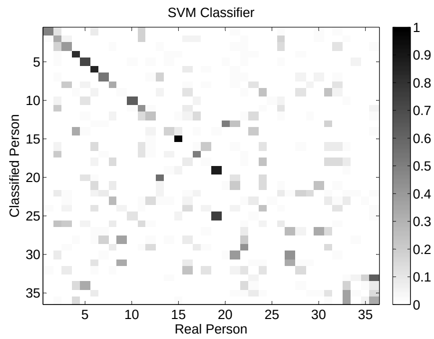
The random chance of correct classification for each picture was 50%. For person-dependent classification with a two-fold cross validation CCR reached 66.20% (64.40%) using the SVM (KNN). The results only slightly were above chance. Possible reasons for the low CCR are discussed later on.

Classification of the fear state in film clips

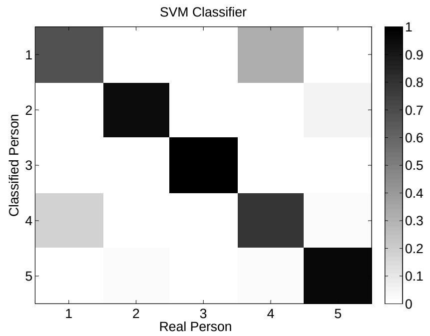
Table 6.1 displays the overall CCR for the SVM and KNN in film-clips using two measurement occasions. Again, the random chance of correct classification was 50%. The CCR fluctuated between 70.30% and 81.90% depending on the classifier and film-clips. In general, the SVM outperformed the KNN. The best CCR was achieved for the SVM, if training data derived from MP 1 *Blairwitch Project* and testing data stemmed from MP 2 *Misery* (81.90%). The CCR was higher than



(a) Confusion matrix within MP 1



(b) Confusion matrix between MP 1 and 2



(c) Confusion matrix between MP 1, 2, and 3

Figure 6.3. Confusion matrix for the correct classification of individuals within- and between measurement points (MPs) using the support vector machine (SVM). Each column represents a real person while each row represents a predicted person. The diagonal shows the probability for a correct classification between 0 and 1 for each individual, false alarms are displayed outside the diagonal. The probability scores can be retrieved from the bar next to the picture. For instance, if a box on the diagonal is black, the person is correctly identified by the classifier. If a box outside the diagonal is black, the classifier identified e.g., individual 19 as individual 25. Data of Figure 6.3(a) and Figure 6.3(b) derived from our experiment. Random chance of correct classification was 2.78% for each individual. Is the respective box on the diagonal darker than 2.78%, the correct classification rate was above chance. Data of Figure 6.3(c) stemmed from another experiment with three measurement points and only five individuals. Random chance of correct classification was 20% for each individual. The results show that the rate of correct classification is higher, if only one measurement occasion and a smaller sample size is used.

the random chance of correct classification. Film-clips achieved better CCRs than pictures.

Further, we were interested in different probabilities for the existence of a fear state over time. Probabilities were calculated using the extension of the SVM classifier and averaged over all participants within each 15 s segment, which explains the stepwise change of the graph in Figure 6.4. The continuous subjective ratings were also averaged across all participants but continuously across the whole time period. Figure 6.4(a) exemplarily depicts the probabilities for the baseline and *The Shining* film clip (MP 1), Figure 6.4(b) is equivalent for the *Misery* film clip on MP 2. First, Figure 6.4 shows that the continuous subjective fear ratings were higher during the fear film clips than during the baseline film clips. Second, the probability for the classification of fear was below 40% for the baseline film clips in MP 1 and MP 2, but above 60% for *The Shining* and *Misery*. This observation stands in line with the results of the rating dial. Third, while the rating dial and classification rates seemed to be more homogeneous during *The Shining*, there were more up- and downturns during *Misery*. Further, the two amplitudes of the rating dial in *Misery* were also visible in peaks of the classification probabilities.

Table 6.1.

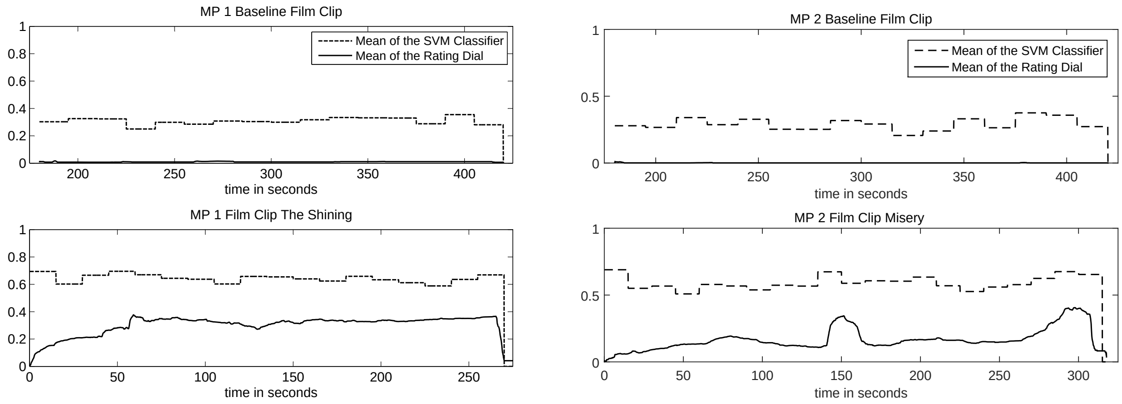
Overall Correct Classification Rate for a Person-Dependent Classification of the Fear State in Film Clips Using two Measurement Occasions. Random Chance of Correct Classification was 50% for Each Film Clip.

	The Shining MP 1		Blairwitch Project MP 1	
	KNN	SVM	KNN	SVM
Misery MP 2	73.40%	74.00%	77.50%	81.90%
Scream MP 2	71.30%	73.30%	70.30%	71.50%

Note. MP = measurement point; KNN = k-nearest neighbors; SVM = support vector machine.

6.4. Discussion

The aims of the present paper were the classification of individuals and the person dependent classification of the emotional state of fear, using physiological data and nonlinear classifiers like the KNN and SVM. We applied pictures and film



(a) Measurement point 1

(b) Measurement point 2

Figure 6.4. Average probabilities for the fear class using an extension of the support vector machine (SVM) and average subjective ratings using the rating dial for the two different measurement points (MPs).

clips to induce fear. Physiological parameters like cardiovascular, electrodermal, and respiratory features were used as underlying modality. Further, continuous subjective ratings with a rating dial were recorded. The contributions of our paper are thus threefold: First, the classification of individuals was included in our analysis. In addition, we had a relatively large sample size with 36 participants. Second, we recorded data on two measurement occasions separated by a time interval of six weeks. Hence, the influence of daily variations (i.e. intraindividual changes) and the reattachment of recording equipment were taken into account in our analysis. Third, we analyzed different probability scores which give a measure of certainty whether the respective segment of film clips belonged to the fear class. We compared these probability scores with continuous subjective ratings. The direct comparison of objectively estimated probability scores and subjective data over the course of a film clip is novel and one of the main contributions of our paper.

In our first research question, we investigated the influence of the number of measurement occasions as well as the number of participants on the classification accuracy of individuals. Our results showed that the CCR decreased, if training and testing data stemmed from different measurement occasions. The result stands in line with other empirical work (e.g., da Silva et al., 2013; Odinaka et al., 2012). A possible explanation for the decline lies in intraindividual variabilities of the

ECG-signal (Schijvenaars, van Herpen, & Kors, 2008). Intraindividual variabilities can be traced back to personal factors like physical activities, drugs, or respiration activity, or to technical factors like differences in the application of the electrodes (Schijvenaars et al., 2008). Further, the CCR increased, when less individuals needed to be classified. The number of overlapping feature vectors are a possible explanation. The higher the number of participants, the higher the probability that feature vectors of different individuals are similar and hence, are overlapping. In that case, it is more difficult to separate individuals from each other and to achieve a correct classification (Bishop, 2006). Consequentially, having more than one measurement occasion and an increasing number of participants impairs the classification accuracy of individuals. Both factors are important for real world applications. For instance, a car, an ATM, or a robot in a nursing home should be able to reliably identify the same individual as well as multiple operators on different days.

In our second research question, we examined the classification accuracy of fear states compared to a neutral, non-emotional state in pictures and film clips. Regarding pictures, our results were only slightly above the random chance of correct classification. On the one hand, this could be due to our two measurement occasions. Previous studies confirmed: Classification accuracy declined if multiple measurement occasions are used for classification (Abdat et al., 2011; R. A. Calvo et al., 2009; Picard et al., 2001). On the other hand, it is possible that the picture time frame of 15 s was not optimally chosen for the analysis of emotion. The onset and duration of an basic emotion should be very quick and short (e.g., Ekman, 2003; Izard, 2007). It is possible that our participants experienced a physiological change at the beginning of the picture, but through the effect of habituation the physiological reaction abated after a few seconds. Unfortunately, studies revealed a high interindividual variability in the duration of emotional episodes, respectively (Kuppens, Oravecz, & Tuerlinckx, 2010; Verduyn, van Mechelen, Tuerlinckx, Meers, & van Coillie, 2009). Further, although, we selected well validated pictures for the induction of fear (Lang et al., 2008), our manipulation check showed that the emotion induction through pictures was probably not intensive enough to yield a high classification accuracy. Compared to the applied pictures, film clips yielded a higher CCR than pictures. This could be due to a more successful emotion induction through additional voice and music elements, as well as moving pictures.

This assumption is supported by our manipulation check, which displayed higher fear ratings for film clips than for pictures.

Overall, the SVM outperformed the KNN and achieved higher CCRs. A reason lies in the type of decision algorithms. The decision of the KNN is solely based on the class label of its nearest neighbours. Especially feature vectors that lie in the border region between two classes are at risk to be misclassified. The SVM, on the other hand, defines a border that separates classes from each other. For each new feature vector, the algorithm checks to which side of the border the feature vector belongs, regardless of the class membership of its neighbours (Bishop, 2006). Thus, depending on the data structure, the SVM might give better classification results than the KNN for our data.

In our third research aim, we compared results of classification analysis with continuous subjective experiences. As emotion induction method, we used film clips, since they do not induce one constant emotional level (Bailenson et al., 2008). Further, we applied an extension of the SVM classifier, which enabled us to calculate probabilities for the fear class. The length of the film clips was divided into 15 s segments, classification was performed for each segment and averaged over all individuals. Our results revealed that the subjective fear rating was higher during the fear film clips than during the neutral baseline film clips. In addition, classification probabilities for the fear class were higher during the fear film clips than during the neutral baseline film clips, as well. The film clips showed different fluctuations in the subjective intensity rating of fear. At the same time, on average, the probabilities for the fear class increased and decreased in accordance with the subjective intensity rating. This observation raises the question if the probabilities could be interpreted as emotion intensity. In the past, intensity levels of emotion were examined by clustering retrospective subjective experiences (Agrawal et al., 2008; Rani et al., 2006). To enable a continuous classification of emotion intensity levels, the extension of the SVM could be a possible solution. Moreover, the duration of the time segments (15 s) can also be reduced and segments can be overlapped to yield a classification that is closer to real-time classification.

Limitations and future directions

Although, our study yielded important implications for the classification of individuals and emotions, there were some limitations to our study. First of all, we used a person-dependent emotion classification approach. Hence, the training and testing data were derived from the same person. For some applications in the real world, it could be necessary that the classifier identifies an emotional state of an unknown person with the help of a large training data base (Chueh et al., 2012). This scenario seems to be a general direction of high interest in emotion classification research (D'Mello & Kory, 2015; Gu et al., 2012; Li et al., 2016), still, some researchers question whether a reliable person-independent classification is even possible (Chueh et al., 2012). The reason for this opinion is a high interindividual variability in physiological signals (Chueh et al., 2012; Jatupaiboon et al., 2015; Li et al., 2016), which was also confirmed by our results. For instance, some individuals were perfectly identified while other individuals had a CCR close to 0%. A possible solution is the combination of person- and emotion classification (Kim & Andre, 2008). For instance, Li et al. (2016) used physiological data and showed that a group based Individual Response Specificity model can achieve higher classification accuracy than a general model. This approach could be a possible solution for the conflict of person-dependent and -independent classification.

Second, we concentrated on physiological signals as underlying modality. Thereby, we did not control for arrhythmia subjects or heart diseases, which could have decreased the recognition rate of individuals (Chiu, Chuang, & Hsu, 2008). Although, participants were explicitly asked to reduce body movements throughout the experiment, motion artifacts can impend classification accuracy. Further, we did not apply feature selection algorithms to our data, since we a priori selected emotion relevant features recording to Gramann and Schandry (2009). The number and kind of features used in classification analysis is very heterogeneous (R. A. Calvo & D'Mello, 2010). Future work should focus on a standardization of physiological features used for classification analysis. A future goal of emotion classification research should be the conjunction of various modalities to optimize the classification and to adjust for the drawbacks of each modality. Studies have shown that the classification rate of physiological data can be improved by adding modalities like facial expression (Abdat et al., 2011; Bailenson et al., 2008).

Third, the degree to which our results can be generalized is limited, since we only induced one emotional state, namely fear. It is even questionable to what extent we only might have induced a higher arousal. Arousal and intesity level. Hence, our results only reflect CCR using two measurement occasions and a low intensity level of fear.

Fourth, we applied well established classification methods like the SVM and

KNN in our study, which are still prominent in the current research (e.g., Jatupaiboon et al., 2015; Li et al., 2016; Wei & Jia, 2016; Wiem & Lachiri, 2017). Due to their context modeling abilities more recent classification methods like Long Short-Term Memory networks (e.g., Wöllmer et al., 2013) or Conditional Random Fields (e.g., Walecki et al., 2017 should be applied to data evolving from two measurement occasions.

The present study provided an important basis for future classification research with psychophysiological data as underlying modality. Our results emphasized the importance of including more than one measurement occasion in the classification analysis. Further, a large sample size was used to yield a more realistic classification setup. Our study presented a possible future approach for analyzing different intensity levels of emotions by using the extension of the SVM classifier. We could show that the probabilities of an emotional class change in accordance with the subjective intensity rating of an emotion. This approach should be extended in future research with more than one emotional state.

References

Abdat, F., Maaoui, C., & Pruski, A. (2011). Bimodal system for emotion recognition from facial expressions and physiological signals using feature-level fusion. In *Proceedings of the UKSim Fifth European Modelling Symposium on Computer Modelling and Simulation* (pp. 24–29). doi: 10.1109/EMS.2011.21

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., & Suri, J. S. (2006). Heart rate variability: A review. *Medical & Biological Engineering & Computing*, *44*(12), 1031–1051. doi: 10.1007/s11517-006-0119-0

Agrawal, P., Liu, C., & Sarkar, N. (2008). Interaction between human and robot: An affect-inspired approach. *Interaction Studies*, *9*(2), 230–257. doi: 10.1075/is.9.2.05agr

Aleknaviciute, J., Tulen, J. H. M., Kamperman, A. M., de Rijke, Y. B., Kooiman, C. G., & Kushner, S. A. (2016). Borderline and cluster C personality disorders manifest distinct physiological responses to psychosocial stress. *Psychoneuroendocrinology*, *72*, 131–138. doi: 10.1016/j.psyneuen.2016.06.010

Al-Nafjan, A., Hosny, M., Al-Wabil, A., & Al-Ohali, Y. (2017). Classification of human emotions from electroencephalogram (EEG) signal using deep neural network. *International Journal of Advanced Computer Science and Applications*, *9*, 419–425. doi: 10.14569/IJACSA.2017.080955

AlZoubi, O., Calvo, R. A., & Stevens, R. H. (2009). Classification of EEG for affect recognition: An adaptive approach. In D. Hutchison et al. (Eds.), *Lecture Notes in Computer Science* (pp. 52–61). doi: 10.1007/978-3-642-10439-8-6

AlZoubi, O., Fossati, D., D'Mello, S., & Calvo, R. A. (2013). Affect detection and classification from the non-stationary physiological data. In *Proceedings of the 12th International Conference on Machine Learning and Applications* (pp. 240–245). doi: 10.1109/ICMLA.2013.49

Attabi, Y., & Dumouchel, P. (2013). Anchor models for emotion recognition from speech. *IEEE Transactions on Affective Computing*, *4*(3), 280–290. doi: 10.1109/T-AFFC.2013.17

Bailenson, J. N., Pontikakis, E. D., Mauss, I. B., Gross, J. J., Jabon, M. E., Hutcherson, C. A. C., ... John, O. (2008). Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human-Computer Studies*, *66*(5), 303–317. doi: 10.1016/j.ijhcs.2007.10.011

Barke, A., Stahl, J., & Kröner-Herwig, B. (2012). Identifying a subset of fear-evoking pictures from the IAPS on the basis of dimensional and categorical ratings for a German sample. *Journal of Behavior Therapy and Experimental Psychiatry*, *43*(1), 565–572. doi: 10.1016/j.jbtep.2011.07.006

Barrett, L. F. (2006a). Are emotions natural kinds? *Perspectives on Psychological Science*, *1*(1), 28–58. doi: 10.1111/j.1745-6916.2006.00003.x

Barrett, L. F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review*, *10*(1), 20–46. doi: 10.1207/s15327957pspr1001{\textunderscore}2

Barrett, L. F. (2009). Variety is the spice of life: A psychological construction approach to understanding variability in emotion. *Cognition & Emotion*, *23*(7), 1284–1306. doi: 10.1080/02699930902985894

Barrett, L. F. (2013). Psychological construction: The darwinian approach to the science of emotion. *Emotion Review*, *5*(4), 379–389. doi: 10.1177/1754073913489753

Barrett, L. F. (2017a). Categories and their role in the science of emotion. *Psychological Inquiry*, *28*(1), 20–26. doi: 10.1080/1047840X.2017.1261581

Barrett, L. F. (2017b). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. doi: 10.1093/scan/nsw154

Barrett, L. F., Quigley, K. S., Bliss-Moreau, E., & Aronson, K. R. (2004). Interoceptive sensitivity and self-reports of emotional experience. *Journal of Personality and Social Psychology*, *87*(5), 684–697. doi: 10.1037/0022-3514.87.5.684

Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*(4), 967–984. doi: 10.1037/0022-3514.74.4.967

Bendat, J. S., & Piersol, A. G. (1990). *Random data: Analysis and measurement procedures.* New York: Wiley.

Ben-Tal, A., Shamailov, S. S., & Paton, J. F. R. (2012). Evaluating the physiological significance of respiratory sinus arrhythmia: looking beyond ventilation-perfusion efficiency. *The Journal of Physiology*, *590*(8), 1989–2008. doi: 10.1113/jphysiol.2011.222422

Bernat, E., Patrick, C. J., Benning, S. D., & Tellegen, A. (2006). Effects of picture content and intensity on affective physiological response. *Psychophysiology*, *43*(1), 93–103. doi: 10.1111/j.1469-8986.2006.00380.x

Berthouze, N., Fushimi, T., Hasegawa, M., Kleinsmith, A., Takenaka, H., & Berthouze, L. (2003). Learning to recognize affective body postures. In *Proceedings of the International Conference Physics and Control.* (pp. 193–198). doi: 10.1109/CIMSA.2003.1227226

Bibbey, A., Carroll, D., Roseboom, T. J., Phillips, A. C., & de Rooij, S. R. (2013). Personality and physiological reactions to acute psychological stress. *International Journal of Psychophysiology*, *90*(1), 28–36. doi: 10.1016/j.ijpsycho.2012.10.018

Biel, L., Pettersson, O., Philipson, L., & Wide, P. (2001). ECG analysis: A new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, *50*(3), 808–812. doi: 10.1109/19.930458

Biopac Systems Inc. (2011). *MP 150 System with AcqKnowledge 4.2.* Goleta, CA: Biopac Systems, Inc.

Bishop, C. M. (2006). *Pattern recognition and machine learning.* New York: Springer.

Boashash, B. (2015). *Time-frequency signal analysis and processing: A comprehensive reference*. London: Academic Press.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

Bonanno, G., & Keltner, D. (2004). The coherence of emotion systems: Comparing "on-line" measures of appraisal and facial expressions, and self-report. *Cognition & Emotion*, *18*(3), 431–444. doi: 10.1080/02699930341000149

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. doi: 10.1037/0033-295x.110.2.203

Boucsein, W. (1988). *Elektrodermale Aktivität: Grundlagen, Methoden und Anwendungen.* Berlin: Springer-Verlag.

Boumbarov, O., Velchev, Y., & Sokolov, S. (2009). ECG personal identification in subspaces using radial basis neural networks. In *Proceedings of the 5th IEEE International Workshop on Intelligent Data Acquisition and Advanced*

Computing Systems, Technology and Applications (pp. 446–451). doi: 10.1109/IDAACS.2009.5342942

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59. doi: 10.1016/0005-7916(94)90063-9

Brcich, R. F., & Iskander, D. R. (2006). Testing for Stationarity in the Frequency Domain Using a Sphericity Statistic. In *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings* (pp. 464–467). IEEE. doi: 10.1109/ICASSP.2006.1660691

Brigham, E. O. (2002). *The Fast Fourier Transform.* New York, NY: Prentice Hall.

Brillinger, D. R. (2001). *Time Series: Data Analysis and Theory* (Vol. 36). New York, NY: Society for Industrial and Applied Mathematics. doi: 10.1137/1.9780898719246

Bulteel, K., Ceulemans, E., Thompson, R. J., Waugh, C. E., Gotlib, I. H., Tuerlinckx, F., & Kuppens, P. (2014). DeCon: A tool to detect emotional concordance in multivariate time series data of emotional responding. *Biological Psychology*, *98*, 29–42. doi: 10.1016/j.biopsycho.2013.10.011

Butler, E. A., Gross, J. J., & Barnard, K. (2014). Testing the effects of suppression and reappraisal on emotional concordance using a multivariate multilevel model. *Biological Psychology*, *98*, 6–18. doi: 10.1016/j.biopsycho.2013.09.003

Cacioppo, J. T., Berntson, G. G., Larsen, J., Poehlmann, K., & Ito, T. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 173–191). New York: Guilford Press.

Cacioppo, J. T., Uchino, B. N., Crites, S. L., Snydersmith, M. A., Smith, G., Berntson, G. G., & Lang, P. J. (1992). Relationship between facial expressiveness and sympathetic activation in emotion - A critical-review, with emphasis on modeling underlying mechanisms and Individual Differences. *Journal of Personality and Social Psychology*, *62*(1), 110–128. doi: 10.1037//0022-3514.62.1.110

Cai, H., & Lin, Y. (2011). Modeling of operators' emotion and task performance in a virtual driving environment. *International Journal of Human-Computer Studies*, *69*(9), 571–586. doi: 10.1016/j.ijhcs.2011.05.003

Calvo, M. G., & Miguel-Tobal, J. J. (1998). The anxiety response: Concordance among components. *Motivation and Emotion*, *22*(3), 211–230. doi: 10.1023/a:1022384022641

Calvo, R. A., Brown, I., & Scheding, S. (2009). Effect of experimental factors on the recognition of affective mental states through physiological measures. In D. Hutchison et al. (Eds.), *Lecture Notes in Computer Science* (pp. 62–70). doi: 10.1007/978-3-642-10439-8_7

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, *1*(1), 18–37. doi: 10.1109/T-AFFC.2010.1

Carroll, D., Lovallo, W. R., & Phillips, A. C. (2009). Are large physiological reactions to acute psychological stress always bad for health? *Social and Personality Psychology Compass*, *3*(5), 725–743. doi: 10.1111/j.1751-9004.2009.00205.x

Ceulemans, E., Kuppens, P., & van Mechelen, I. (2012). Capturing the structure of distinct types of individual differences in the situation-specific experience of emotions: The case of anger. *European Journal of Personality*, *26*(5), 484–495. doi: 10.1002/per.847

Chan, A. D. C., Hamdy, M. M., Badre, A., & Badee, V. (2006). Person identification using electrocardiograms. In *Canadian Conference on Electrical and Computer Engineering* (pp. 1–4). doi: 10.1109/CCECE.2006.277291

Chanel, G., Kierkels, J. J., Soleymani, M., & Pun, T. (2009). Short-term emotion assessment in a recall paradigm. *International Journal of Human-Computer Studies*, *67*(8), 607–627. doi: 10.1016/j.ijhcs.2009.03.005

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1–27. doi: 10.1145/1961189.1961199

Chang, C.-Y., Tsai, J.-S., Wang, C.-J., & Chung, P.-c. (2009). Emotion recognition with consideration of facial expression and physiological signals. In *Proceedings of the 9th IEEE International Symposium on Signal Processing and Information Technology* (pp. 278–283). doi: 10.1109/CIBCB.2009.4925739

Chaplin, T. M. (2014). Gender and emotion expression: A developmental contextual perspective. *Emotion Review*, *7*(1), 14–21. doi: 10.1177/1754073914544408

Chen, S.-W., Chen, H.-C., & Chan, H.-L. (2006). A real-time 5QRS6 detection method based on moving-averaging incorporating with wavelet denoising. *Computer Methods and Programs in Biomedicine*, *82*(3), 187–195. doi: 10.1016/j.cmpb.2005.11.012

Cheng, K.-S., Chen, Y.-S., & Wang, T. (2012). Physiological parameters assessment for emotion recognition. In *Proceedings of the International Conference on Biomedical and Health Informatics* (pp. 995–998). doi: 10.1109/IECBES.2012.6498118

Chiu, C.-C., Chuang, C.-M., & Hsu, C.-Y. (2008). A novel personal identity verification approach using a discrete wavelet transform of the ECG signal. In *International Conference on Multimedia and Ubiquitous Engineering* (pp. 201–206). doi: 10.1109/MUE.2008.67

Christie, I. C., & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *International Journal of Psychophysiology*, *51*(2), 143–153. doi: 10.1016/j.ijpsycho.2003.08.002

Chueh, T.-H., Chen, T.-B., Lu, H. H.-S., Ju, S.-S., Tao, T.-H., & Shaw, J.-H. (2012). Statistical predicition of emotional states by physiological signals with MANOVA and machine learning. *International Journal of Pattern Recognition and Artificial Intelligence*, *26*(4). doi: 10.1142/S0218001412500085

Clark-Polner, E., Johnson, T. D., & Barrett, L. F. (2017). Multivoxel pattern analysis does not provide evidence to support the existence of basic emotions. *Cerebral Cortex*, *27*(3), 1944–1948. doi: 10.1093/cercor/bhw028

Clifford, G. D. (2002). *Signal processing methods for heart rate variability* (Doctoral dissertation, University of Oxford, Oxford, Great Britain). Retrieved from `http://www.ibme.ox.ac.uk/research/biomedical-signal-processing-instrumentation/prof-l-tarassenko/publications/pdf/gdcliffordthesis.pdf`

Clore, G. L., & Ortony, A. (2013). Psychological construction in the OCC model of emotion. *Emotion Review*, *5*(4), 335–343. doi: 10.1177/1754073913489751

Coan, J. A. (2010). Emergent ghosts of the emotion machine. *Emotion Review*, *2*(3), 274–285. doi: 10.1177/1754073910361978

Codispoti, M., & de Cesarei, A. (2007). Arousal and attention: Picture size and emotional reactions. *Psychophysiology*, *44*(5), 680–686. doi: 10.1111/j.1469-8986.2007.00545.x

Cohen, M. X. (2008). Assessing transient cross-frequency coupling in EEG data. *Journal of Neuroscience Methods*, *168*(2), 494–499. doi: 10.1016/j.jneumeth

.2007.10.012

Costa, T., Rognoni, E., & Galati, D. (2006). EEG phase synchronization during emotional response to positive and negative film stimuli. *Neuroscience Letters*, *406*(3), 159–164. doi: 10.1016/j.neulet.2006.06.039

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27. doi: 10.1109/TIT.1967.1053964

Cunningham, W. A., Dunfield, K. A., & Stillman, P. E. (2013). Emotional states from affective dynamics. *Emotion Review*, *5*(4), 344–355. doi: 10.1177/1754073913489749

da Silva, H. P., Fred, A., Lourenco, A., & Jain, A. K. (2013). Finger ECG signal for user authentication: Usability and performance. In *Proceedings of the 6th International Conference on Biometrics: Theory, Applications and Systems* (pp. 1–8). doi: 10.1109/BTAS.2013.6712689

Damasio, A. R., Grabowski, T. J., Bechara, A., Damasio, H., Ponto, L. L., Parvizi, J., & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*(10), 1049–1056. doi: 10.1038/79871

Dan-Glauser, E. S., & Gross, J. J. (2011). The temporal dynamics of two response-focused forms of emotion regulation: Experiential, expressive, and autonomic consequences. *Psychophysiology*, *48*(9), 1309–1322. doi: 10.1111/j.1469-8986.2011.01191.x

Dan-Glauser, E. S., & Gross, J. J. (2013). Emotion regulation and emotion coherence: Evidence for strategy-specific effects. *Emotion*, *13*(5), 832–842. doi: 10.1037/a0032672

Darwin, C. (1872/1965). *The expression of emotion in man and animals.* Chicago: University of Chicago Press.

Davidson, R. J. (1978). Specificity and patterning in biobehavioral systems: Implications for behavior change. *American Psychologist*, *33*(5), 430–436. doi: 10.1037/0003-066X.33.5.430

Davidson, R. J. (1992). Prolegomenon to the structure of emotion: Gleanings from neuropsychology. *Cognition & Emotion*, *6*(3), 245–268. doi: 10.1080/02699939208411071

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of* *psychophysiology* (pp. 200–223). Cambridge: Cambridge University Press.

Demaree, H. A., Schmeichel, B. J., Robinson, J. L., Pu, J., Everhart, & Berntson, G. G. (2006). Up- and down-regulating facial disgust: Affective, vagal, sympathetic, and respiratory consequences. *Biological Psychology*, *71*(1), 90–99. doi: 10.1016/j.biopsycho.2005.02.006

DeSteno, D., Petty, R. E., Wegener, D. T., & Rucker, D. D. (2000). Beyond valence in the perception of likelihood: The role of emotion specificity. *Journal of Personality and Social Psychology*, *78*(3), 397–416. doi: 10.1037/0022-3514.78.3.397

Dixon-Gordon, K. L., Yiu, A., & Chapman, A. L. (2013). Borderline personality features and emotional reactivity: The mediating role of interpersonal vulnerabilities. *Journal of Behavior Therapy and Experimental Psychiatry*, *44*(2), 271–278. doi: 10.1016/j.jbtep.2012.12.001

D'Mello, S. K. (2013). A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology*, *105*(4), 1082–1099. doi: 10.1037/a0032674

D'Mello, S. K., Craig, S. D., Witherspoon, A., McDaniel, B., & Graesser, A. (2008). Automatic detection of learner's affect from conversational cues. *User Modeling and User-Adapted Interaction*, *18*(1-2), 45–80. doi: 10.1007/s11257-007-9037-6

D'Mello, S. K., & Kory, J. (2015). A review and meta-analysis of multimodal affect detection systems. *Acm Computing Surveys*, *47*(3). doi: 10.1145/2682899

Dougherty, J., Kohavi, R., & Sahami, M. (1995). Supervised and unsupervised discretization of continuous features. In A. Prieditis & S. Russell (Eds.), *Machine learning: Proceedings of the 12th internationa conference* (pp. 194–202). San Francisco, CA: Morgan Kaufmann Publishers.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods* (Second ed.). Oxford: Oxford University Press.

Ehlers, A., Suendermann, O., Boellinghaus, I., Vossbeck-Elsebusch, A., Gamer, M., Briddon, E., … Glucksman, E. (2010). Heart rate responses to standardized trauma-related pictures in acute posttraumatic stress disorder. *International Journal of Psychophysiology*, *78*(1), 27–34. doi: 10.1016/j.ijpsycho.2010.04.009

Ekman, P. (1984). Expression and the nature of emotion. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 319–343). Hillsdale, NJ:

Lawrence Erlbaum.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3), 169–200. doi: 10.1080/02699939208411068

Ekman, P. (2003). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life* (1st ed.). New York: Times Books.

Ekman, P. (2016). What scientists who study emotion agree about. *Perspectives on Psychological Science*, *11*(1), 31–34. doi: 10.1177/1745691615596992

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3*(4), 364–370. doi: 10.1177/1754073911410740

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous-system activity distinguishes among emotions. *Science*, *221*(4616), 1208–1210. doi: 10.1126/science.6612338

Esslen, M., Pascual-Marqui, R. D., Hell, D., Kochi, K., & Lehmann, D. (2004). Brain areas and time course of emotional processing. *Neuroimage*, *21*(4), 1189–1203. doi: 10.1016/j.neuroimage.2003.10.001

Evers, C., Hopp, H., Gross, J. J., Fischer, A. H., Manstead, Antony S. R., & Mauss, I. B. (2014). Emotion response coherence: A dual-process perspective. *Biological Psychology*, *98*, 43–49. doi: 10.1016/j.biopsycho.2013.11.003

Fernandez, C., Pascual, J. C., Soler, J., Elices, M., Portella, M. J., & Fernandez-Abascal, E. (2012). Physiological responses induced by emotion-eliciting films. *Applied Psychophysiology and Biofeedback*, *37*(2), 73–79. doi: 10.1007/s10484-012-9180-7

Fix, E., & Hodges, J. L.,Jr. (1989). Discriminatory analysis. Nonparametric discrimination: Consistency properties. *International Statistical Review*, *57*(3), 238–247. Retrieved from http://www.jstor.org/stable/1403797

Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050–1057. doi: 10.1111/j.1467-9280.2007.02024.x

Fratini, A., Sansone, M., Bifulco, P., Romano, M., Pepino, A., Cesarelli, M., & D'Addio, G. (2013). Individual identification using electrocardiogram morphology. In *Proceedings of the International Symposium on Medical Measurements and Applications* (pp. 107–110). doi: 10.1109/MeMeA.2013.6549716

Fredrickson, B. L., & Levenson, R. W. (1998). Positive emotions speed recovery from the cardiovascular sequelae of negative emotions. *Cognition & Emotion*, *12*(2), 191–220. doi: 10.1080/026999398379718

Friedman, B. H., & Kreibig, S. D. (2010). The biopsychology of emotion: Current theoretical, empirical, and methodological perspectives. *Biological Psychology*, *84*(3), 381–382. doi: 10.1016/j.biopsycho.2010.02.001

Friedman, B. H., Stephens, C. L., & Thayer, J. F. (2014). Redundancy analysis of autonomic and self-reported responses to induced emotions. *Biological Psychology*, *98*, 19–28. doi: 10.1016/j.biopsycho.2013.12.006

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474–480. doi: 10.1016/j.tics.2005.08.011

Garcia, A. J., Koschnitzky, J. E., Dashevskiy, T., & Ramirez, J.-M. (2013). Cardiorespiratory coupling in health and disease. *Autonomic Neuroscience*, *175*(1-2), 26–37. doi: 10.1016/j.autneu.2013.02.006

Gentsch, K., Grandjean, D., & Scherer, K. R. (2014). Coherence explored between emotion components: Evidence from event-related potentials and facial electromyography. *Biological Psychology*, *98*, 70–81. doi: 10.1016/j.biopsycho.2013.11.007

Gilbert, P. (2013). *Brief user's guide: Dynamic systems estimation (DSE).* Retrieved from `http://cran.r-project.org/web/packages/dse/vignettes/dse-guide.pdf`

Gilbert, P. D. (1993). *State space and ARMA models: An overview of the Equivalene.* (Tech. Rep. No. 94-3). Bank of Canada.

Goodkind, M. S., Sollberger, M., Gyurak, A., Rosen, H. J., Rankin, K. P., Miller, B., & Levenson, R. (2012). Tracking emotional valence: The role of the orbitofrontal cortex. *Human Brain Mapping*, *33*(4), 753–762. doi: 10.1002/hbm.21251

Gouizi, K., Bereksi-Reguig, F., & Maaoui, C. (2011). Emotion recognition from physiological signals. *Journal of Medical Engineering & Technology*, *35*(6-7), 300–307. doi: 10.3109/03091902.2011.601784

Gramann, K., & Schandry, R. (2009). *Psychophysiologie: Körperliche Indikatoren psychischen Geschehens* (4th ed.). Weinheim, Basel: Beltz, PVU.

Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal-driven response synchronization. *Consciousness and Cognition*, *17*(2), 484–495. doi: 10.1016/j.concog.2008.03.019

Grewal, M. S., & Andrews, A. P. (2001). *Kalman filtering: Theory and practice* *using MATLAB®* (Second ed.). Cambridge: Cambridge University Press.

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(5/6), 561–566. doi: 10.5194/npg-11-561-2004

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*(1), 224–237. doi: 10.1037//0022-3514.74.1.224

Gross, J. J. (2010). The future's so bright, I gotta wear shades. *Emotion Review*, *2*(3), 212–216. doi: 10.1177/1754073910361982

Gross, J. J., & Levenson, R. W. (1993). Emotional suppression: Physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, *64*(6), 970–986. doi: 10.1037//0022-3514.64.6.970

Gross, J. J., & Levenson, R. W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, *106*(1), 95–103. doi: 10.1037/0021-843x.106.1.95

Gu, Y., Tan, S. L., Wong, K. J., Ho, M. H. R., & Qu, L. (2009). Using GA-based feature selection for emotion recognition from physiological signals. In *Proceedings of the 9th IEEE International Symposium on Signal Processing and Information Technology* (pp. 1–4). doi: 10.1109/ISPACS.2009.4806747

Gu, Y., Wong, K.-J., & Tan, S.-L. (2012). Analysis of physiological responses from multiple subjects for emotion recognition. In *Proceedings of the 14th International Conference on e-Health Networking, Applications and Services* (pp. 178–183). doi: 10.1109/HealthCom.2012.6379388

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN Model-based approach in classification. In R. Meersman, Z. Tari, & Schmidt, Douglas, C. (Eds.), *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE* (Vol. 2888, pp. 986–996). doi: 10.1007/978-3-540-39964-3_62

Gyurak, A., Gross, J. J., & Etkin, A. (2011). Explicit and implicit emotion regulation: A dual-process framework. *Cognition & Emotion*, *25*(3), 400–412. doi: 10.1080/02699931.2010.544160

Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion recognition using bio-sensors: First steps towards an automatic system. In T. Kanade et al. (Eds.), *Affective Dialogue Systems* (Vol. 3068, pp. 36–48). doi: 10.1007/978-3-540-24842_4

Hamaker, E. L., Dolan, C. V., & Molenaar, P. C. (2005). Statistical modeling of the individual: Rationale and application of multivariate stationary time series analysis. *Multivariate Behavioral Research*, *40*(2), 207–233. doi: 10.1207/s15327906mbr4002-3

Hastings, P. D., Nuselovici, J. N., Klimes-Dougan, B., Kendziora, K. T., Usher, B. A., Ho, M.-h. R., & Zahn-Waxler, C. (2009). Dysregulated coherence of subjective and cardiac emotional activation in adolescents with internalizing and externalizing problems. *Journal of Child Psychology and Psychiatry*, *50*(11), 1348–1356. doi: 10.1111/j.1469-7610.2009.02159.x

Haykin, S. S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Upper Saddle River, New Jersey: Prentice Hall.

Herbert, B. M., Pollatos, O., & Schandry, R. (2007). Interoceptive sensitivity and emotion processing: An EEG study. *International Journal of Psychophysiology*, *65*(3), 214–227. doi: 10.1016/j.ijpsycho.2007.04.007

Herpertz, S. C., & Koetting, K. (2005). Startle response in inpatients with borderline personality disorder vs. healthy controls. *Journal of Neural Transmission*, *112*(8), 1097–1106. doi: 10.1007/s00702-004-0249-1

Herring, D. R., Burleson, M. H., Roberts, N. A., & Devine, M. J. (2011). Coherent with laughter: Subjective experience, behavior, and physiological responses during amusement and joy. *International Journal of Psychophysiology*, *79*(2), 211–218. doi: 10.1016/j.ijpsycho.2010.10.007

Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition & Emotion*, *19*(7), 1095–1109. doi: 10.1080/02699930541000084

Hinz, A., Hueber, B., Schreinicke, G., & Seibt, R. (2002). Temporal stability of psychophysiological response patterns: concepts and statistical tools. *International Journal of Psychophysiology*, *44*(1), 57–65. doi: 10.1016/S0167-8760(01)00191-X

Hlawatsch, F., & Auger, F. (Eds.). (2013). *Time-frequency analysis.* New York: Wiley-Liss.

Hodgson, R., & Rachman, S. (1974). II. Desynchrony in measures of fear. *Behaviour Research and Therapy*, *12*(4), 319–326. doi: 10.1016/0005-7967(74)90006-0

Hollenstein, T., & Lanteigne, D. (2014). Models and methods of emotional concordance. *Biological Psychology*, *98*, 1–5. doi: 10.1016/j.biopsycho.2013.12.012

Hornak, J. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain*, *126*(7), 1691–1712. doi: 10.1093/brain/awg168

Hot, P., & Sequeira, H. (2013). Time course of brain activation elicited by basic emotions. *NeuroReport*, *24*(16), 898–902. doi: 10.1097/WNR.0000000000000016

Hsieh, F., Ferrer, E., Chen, S., Mauss, I. B., John, O., & Gross, J. J. (2011). A network approach for evaluating coherence in multivariate systems: An application to psychophysiological emotion data. *Psychometrika*, *76*(1), 124–152. doi: 10.1007/s11336-010-9194-0

Hudlicka, E. (2003). To feel or not to feel: The role of affect in human?computer interaction. *International Journal of Human-Computer Studies*, *59*, 1–32. doi: 10.1016/S1071-5819(03)00047-8

Hulsman, R. L., Smets, E. M. A., Karemaker, J. M., & de Haes, H. J. C. J. M. (2011). The psychophysiology of medical communication. Linking two worlds of research. *Patient Education and Counseling*, *84*(3), 420–427. doi: 10.1016/j.pec.2011.05.008

Izard, C. E. (1977). *Human emotions.* New York: Plenum Press.

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science*, *2*(3), 260–280. doi: 10.1111/j.1745-6916.2007.00044.x

Izard, C. E. (2011). Forms and functions of emotions: Matters of emotion-cognition interactions. *Emotion Review*, *3*(4), 371–378. doi: 10.1177/1754073911410737

Izard, C. E., Dougherty, F. E., Bloxom, B. M., & Kotsch, W. E. (1974). *The Differential Emotions Scale: A method of measuring the subjective experience of discrete emotions.* Unpublished manuscript, Departement of Psychology, University Vanderbilt, Nashville, TN.

James, W. (1884). What is an emotion? *Mind*(34), 188–205. doi: 10.1093/mind/os-IX.34.188

Jang, E.-H., Park, B.-J., Kim, S.-H., & Sohn, J.-H. (2012). Emotion classification based on physiological signals induced by negative emotions: Discriminantion of negative emotions by machine learning algorithm. In *Proceedings of 9th IEEE International Conference on Networking, Sensing and Control* (pp. 283–288). doi: 10.1109/ICNSC.2012.6204931

Jatupaiboon, N., Pan-Ngum, S., & Israsena, P. (2015). Subject-dependent and subject-independent emotion classification using unimodal and multimodal physiological signals. *Journal of Medical Imaging and Health Informatics*, *5*(5), 1020–1027. doi: 10.1166/jmihi.2015.1490

Jerritta, S., Murugappan, M., Nagarajan, R., & Wan, K. (2011). Physiological signals based human emotion recognition: A review. In *Proceedings of the 7th International Colloquium on Signal Processing and its Applications* (pp. 410–415). doi: 10.1109/CSPA.2011.5759912

Kassam, K. S., Markey, A. R., Cherkassky, V. L., Loewenstein, G., & Just, M. A. (2013). Identifying emotions on the basis of neural activation. *PLoS ONE*, *8*(6), e66032. doi: 10.1371/journal.pone.0066032

Keissar, K., Davrath, L. R., & Akselrod, S. (2009). Coherence analysis between respiration and heart rate variability using continuous wavelet transform. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *367*(1892), 1393–1406. doi: 10.1098/rsta.2008.0273

Kelava, A., Muma, M., Deja, M., Dagdagan, J. Y., & Zoubir, A. M. (2015). A new approach for the quantification of synchrony of multivariate non-stationary psychophysiological variables during emotion eliciting stimuli. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01507

Kettunen, J., Ravaja, N., Näätänen, P., Keskivaara, P., & Keltikangas-Järvinen, L. (1998). The synchronization of electrodermal activity and heart rate and its relationship to energetic arousal: a time series approach. *Biological Psychology*, *48*(3), 209–225. doi: 10.1016/S0301-0511(98)00017-9

Kim, J., & Andre, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(12), 2067–2083. doi: 10.1109/TPAMI.2008.26.

Kleinsmith, A., Bianchi-Berthouze, N., & Steed, A. (2011). Automatic recognition of non-acted affective postures. *IEEE Trans. Syst., Man, Cybern. B (IEEE Transactions on Systems, Man, and Cybernetics, Part B)*, *41*(4), 1027–1038. doi: 10.1109/TSMCB.2010.2103557

Kolodyazhniy, V., Kreibig, S. D., Gross, J. J., Roth, W. T., & Wilhelm, F. H. (2011). An affective computing approach to physiological emotion specificity: Toward subject-independent and stimulus-independent classification of film-induced emotions. *Psychophysiology*, *48*(7), 908–922. doi: 10.1111/j.1469-8986.2010.01170.x

Kortelainen, J., Tiinanen, S., Huang, X., Li, X., Laukka, S., Pietikainen, M., & Seppanen, T. (2012). Multimodal emotion recognition by combining physiological signals and facial expressions: A preliminary study. In *Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5238–5241). doi: 10.1109/EMBC.2012.6347175

Kragel, P. A., & LaBar, K. S. (2013). Multivariate pattern classification reveals autonomic and experiential representations of discrete emotions. *Emotion*, *13*(4), 681–690. doi: 10.1037/a0031820

Kragel, P. A., & LaBar, K. S. (2015). Multivariate neural biomarkers of emotional states are categorically distinct. *Social Cognitive and Affective Neuroscience*, *10*(11), 1437–1448. doi: 10.1093/scan/nsv032

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: A review. *Biological Psychology*, *84*(3), 394–421. doi: 10.1016/j.biopsycho.2010.03.010

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology*, *44*(5), 787–806. doi: 10.1111/j.1469-8986.2007.00550.x

Kring, A. M., & Gordon, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality and Social Psychology*, *74*(3), 686–703. doi: 10.1037/0022-3514.74.3.686

Kristjansson, S. D., Kircher, J. C., & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: An introduction to growth curve modeling. *Psychophysiology*, *44*(5), 728–736. doi: 10.1111/j.1469-8986.2007.00544.x

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology*, *51*(4), 336–353. doi: 10.1037/1196-1961.51.4.336

Kukolja, D., Popovic, S., Horvat, M., Kovac, B., & Cosic, K. (2014). Comparative analysis of emotion estimation methods based on physiological measurements for real-time applications. *International Journal of Human-Computer Studies*, *72*(10-11), 717–727. doi: 10.1016/j.ijhcs.2014.05.006

Kuppens, P., Champagne, D., & Tuerlinckx, F. (2012). The Dynamic Interplay between Appraisal and Core Affect in Daily Life. *Frontiers in Psychology*, *3*. doi: 10.3389/fpsyg.2012.00380

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, *99*(6), 1042–1060. doi: 10.1037/a0020962

Kuppens, P., Stouten, J., & Mesquita, B. (2009). Individual differences in emotion components and dynamics: Introduction to the Special Issue. *Cognition & Emotion*, *23*(7), 1249–1258. doi: 10.1080/02699930902985605

Kuppens, P., van Mechelen, I., Smits, D. J. M., de Boeck, P., & Ceulemans, E. (2007). Individual differences in patterns of appraisal and anger experience. *Cognition & Emotion*, *21*(4), 689–713. doi: 10.1080/02699930600859219

Lacey, J. I., & Lacey, B. C. (1958). Verification and extension of the principle of autonomic response-stereotypy. *The American Journal of Psychology*, *71*(1), 50. doi: 10.2307/1419197

Lander, P., & Berbari, E. J. (1997). Time-frequency plane Wiener filtering of the high-resolution ECG: development and application. *IEEE Transactions on Biomedical Engineering*, *44*(4), 256–265. doi: 10.1109/10.563295

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual.* (Technical Report A-8). Gainesville, FL.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures - affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. doi: 10.1111/j.1469-8986.1993.tb03352.x

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776–785. doi: 10.1111/1469-8986.00078

LaVaque, T. J. (2003). Neurofeedback, Neurotherapy, and Quantitative EEG. In D. P. Moss, A. McGrady, & T. Davies (Eds.), *Handbook of mind-body medicine for primary care* (pp. 123–136). Thousand Oaks, CA: Sage Publications.

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, *46*(8), 819–834. doi: 10.1037/0003-066X.46.8.819

Lazarus, R. S., Speisman, J. C., & Mordkoff, A. M. (1963). The relationship between autonomic indicators of psychological stress: Heart rate and skin conductance. *Psychosomatic Medicine*, *25*(1), 19–30.

LeDoux, J. E. (1989). Cognitive-emotional interactions in the brain. *Cognition &*

Emotion, *3*(4), 267–289. doi: 10.1080/02699938908412709

LeDoux, J. E. (2000). Emotion Circuits in the Brain. *Annual Review of Neuroscience*, *23*(1), 155–184. doi: 10.1146/annurev.neuro.23.1.155

Lee, C. M., & Narayanan, S. S. (2005). Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, *13*(2), 293–303. doi: 10.1109/TSA.2004.838534

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*(5), 834–855. doi: 10.1037/a0024244

Leng, H., Lin, Y., & Zanzi, L. A. (2007). An experimental study on physiological parameters toward driver emotion recognition. In M. J. Dainoff (Ed.), *Proceedings of the International Conference of Ergonomics and Health Aspects of Work with Computers held as part of the International Conference on Human-Computer Interaction* (pp. 237–246). doi: 10.1007/978-3-540-73333-1_30

Lerner, J. S., Dahl, R. E., Hariri, A. R., & Taylor, S. E. (2007). Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses. *Biological Psychiatry*, *61*(2), 253–260. doi: 10.1016/j.biopsych.2006.08.016

Levenson, R. W. (1992). Autonomic nervous-system differences among emotions. *Psychological Science*, *3*(1), 23–27. doi: 10.1111/j.1467-9280.1992.tb00251.x

Levenson, R. W. (1994). Human emotion: A functional view. In P. Ekman (Ed.), *The nature of emotion* (pp. 123–126). New York: Oxford Univ. Press.

Levenson, R. W. (2003). Autonomic specificity and emotion. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 212–224). Oxford, New York: Oxford University Press.

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review*, *3*(4), 379–386. doi: 10.1177/1754073911410743

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous-system activity. *Psychophysiology*, *27*(4), 363–384. doi: 10.1111/j.1469-8986.1990.tb02330.x

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences*, *28*(02). doi: 10.1017/S0140525X0500004X

Li, C., Xu, C., & Feng, Z. (2016). Analysis of physiological for emotion recognition with the IRS model. *Neurocomputing*, *178*, 103–111. doi: 10.1016/j.neucom

.2015.07.112

Lim, C. L., Rennie, C., Barry, R. J., Bahramali, H., Lazzaro, I., Manor, B., & Gordon, E. (1997). Decomposing skin conductance into tonic and phasic components. *International Journal of Psychophysiology*, *25*(2), 97–109. doi: 10.1016/S0167-8760(96)00713-1

Lindquist, K. A. (2013). Emotions emerge from more basic psychological ingredients: A modern psychological constructionist model. *Emotion Review*, *5*(4), 356–368. doi: 10.1177/1754073913489750

Lindquist, K. A., Siegel, E. H., Quigley, K. S., & Barrett, L. F. (2013). The hundred-year emotion war: Are emotions natural kinds or psychological constructions? Comment on Lench, Flores, and Bench (2011). *Psychological Bulletin*, *139*(1), 255–263. doi: 10.1037/a0029038

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*(03), 121–143. doi: 10.1017/S0140525X11000446

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: The Guilford Press.

Lisetti, C. L., & Nasoz, F. (2004). Using noninvasive wearable computers to recognize human emotions from physiological signals. *EURASIP Journal on Advances in Signal Processing*, *2004*(11), 1672–1687. doi: 10.1155/S1110865704406192

Liu, C., Conn, K., Sarkar, N., & Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics*, *24*(4), 883–896. doi: 10.1109/TRO.2008.2001362

Lourenco, A., Silva, H., & Fred, A. (2012). ECG-based biometrics: A real time classification approach. In I. Santamaría, J. Arenas-García, G. Camps-Valls, D. Erdogmus, F. Pérez-Cruz, & J. Larsen (Eds.), *IEEE International Workshop on Machine Learning for Signal Processing* (pp. 1–6). doi: 10.1109/MLSP.2012.6349735

Lykken, D., & Tellegen, A. (1996). Happiness is a stochastic phenomenon. *Psychological Science*, *7*(3), 186–189. doi: 10.1111/j.1467-9280.1996.tb00355.x

Lyusin, D., & Ovsyannikova, V. (2016). Measuring two aspects of emotion recognition ability: Accuracy vs. sensitivity. *Learning and Individual Differences*, *52*, 129–136. doi: 10.1016/j.lindif.2015.04.010

Maaoui, C., & Pruski, A. (2008). A comparative study of SVM kernel applied to emotion recognition from physiological signals. In *Proceedings of the International Conference on Computational Intelligence and Security* (pp. 1–6). doi: 10.1109/SSD.2008.4632891

MacLean, P. D. (1990). *The triune brain in evolution: Role in paleocerebral functions*. New York: Plenum P.

Marci, C. D., & Orr, S. P. (2006). The effect of emotional distance on psychophysiologic concordance and perceived empathy between patient and interviewer. *Applied Psychophysiology and Biofeedback*, *31*(2), 115–128. doi: 10.1007/s10484-006-9008-4

Marple, S. (1987). *Digital Spectral Analysis: With Applications*. Englewood Cliffs, NJ: PrenticeHall.

Marwitz, M., & Stemmler, G. (1998). On the status of individual response specificity. *Psychophysiology*, *35*(1), 1–15.

Marzillier, S. L., & Davey, G. C. (2005). Anxiety and disgust: Evidence for a unidirectional relationship. *Cognition & Emotion*, *19*(5), 729–750. doi: 10.1080/02699930441000436

MathWorks. (2011). *Matlab R2011b*. Natick, MA: The Math Works, Inc.

Matta, R., Lau, J. K. H., Agrafioti, F., & Hatzinakos, D. (2011). Real-time continuous identification system using ECG signals. In *24th Canadian Conference on Electrical and Computer Engineering* (pp. 001313–001316). doi: 10.1109/CCECE.2011.6030676

Matz, G., & Hlawatsch, F. (2000). Time-frequency coherence analysis of nonstationary random processes. In *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing* (pp. 554–558). IEEE. doi: 10.1109/SSAP.2000.870186

Mauss, I. B., Levenson, R. W., McCarter, L., Wilhelm, F. H., & Gross, J. J. (2005). The tie that binds? Coherence among emotion experience, behavior, and physiology. *Emotion*, *5*(2), 175–190. doi: 10.1037/1528-3542.5.2.175

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, *23*(2), 209–237. doi: 10.1080/02699930802204677

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 & Emotion*, *18*(5), 631–662. doi: 10.1080/02699930341000112

McAssey, M. P., Helm, J., Hsieh, F., Sbarra, D. A., & Ferrer, E. (2013). Methodological Advances for Detecting Physiological Synchrony During Dyadic Interactions. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *9*(2), 41–53. doi: 10.1027/1614-2241/a000053

Merten, J., & Krause, R. (1993). *DAS (Differentielle Affekt Skala).* Unpublished manuscript, Department of Psychology, Saarland University, Saarbrücken, Germany.

Meuret, A. E., Rosenfield, D., Wilhelm, F. H., Zhou, E., Conrad, A., Ritz, T., & Roth, W. T. (2011). Do Unexpected Panic Attacks Occur Spontaneously? *Neurodevelopment and Depression*, *70*(10), 985–991. doi: 10.1016/j.biopsych.2011.05.027

Mikels, J. A., Fredrickson, B. L., Larkin, G. R., Lindberg, C. M., Maglio, S. J., & Reuter-Lorenz, P. A. (2005). Emotional category data on images from the international affective picture system. *Behavior Research Methods*, *37*(4), 626–630. doi: 10.3758/BF03192732

Miskovic, V., & Schmidt, L. A. (2010). Cross-regional cortical synchronization during affective image viewing. *Brain Research*, *1362*, 102–111. doi: 10.1016/j.brainres.2010.09.102

Molenaar, P. C. M. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. *Measurement: Interdisciplinary Research & Perspective*, *2*(4), 201–218. doi: 10.1207/s15366359mea0204{\textunderscore}1

Moors, A. (2009). Theories of emotion causation: A review. *Cognition & Emotion*, *23*(4), 625–662. doi: 10.1080/02699930802645739

Muma, M., Iskander, D., & Collins, M. (2010). The role of cardiopulmonary signals in the dynamics of the eye's wavefront aberrations. *IEEE Transactions on Biomedical Engineering*, *57*(2), 373–383. doi: 10.1109/tbme.2009.2032531

Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(3), 207–233. doi: 10.3758/cabn.3.3.207

Nesse, R. M., Curtis, G. C., Thyer, B. A., McCann, D. S., Huber-Smith, M. J., & Knopf, R. F. (1985). Endocrine and cardiovascular responses during phobic anxiety. *Psychosomatic Medicine*, *47*(4), 320–332.

Nesselroade, J. R., & Molenaar, P. C. M. (1999). Pooling lagged covariance structures based on short, multivariate time-series for dynamic facto ranalysis. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 224–251). Newbury Park: Sage Publ.

Nezlek, J. B., Vansteelandt, K., van Mechelen, I., & Kuppens, P. (2008). Appraisal-emotion relationships in daily life. *Emotion*, *8*(1), 145–150. doi: 10.1037/1528-3542.8.1.145

Niskanen, J.-P., Tarvainen, M. P., Ranta-Aho, P. O., & Karjalainen, P. A. (2004). Software for advanced HRV analysis. *Computer Methods and Programs in Biomedicine*, *76*(1), 73–81. doi: 10.1016/j.cmpb.2004.03.004

Novak, P., & Novak, V. (1993). Time-frequency mapping of the heart rate, blood pressure and respiratory signals. *Medical & Biological Engineering & Computing*, *31*(2), 103–110. doi: 10.1007/BF02446667

Oatley, K., & Johnson-laird, P. N. (1987). Towards a Cognitive Theory of Emotions. *Cognition & Emotion*, *1*(1), 29–50. doi: 10.1080/02699938708408362

Odinaka, I., Lai, P.-H., Kaplan, A. D., O'Sullivan, J. A., Sirevaag, E. J., & Rohrbaugh, J. W. (2012). ECG biometric recognition: A comparative analysis. *IEEE Transactions on Information Forensics and Security*, *7*(6), 1812–1824. doi: 10.1109/TIFS.2012.2215324

Oliveira-Silva, P., & Gonçalves, O. F. (2011). Responding empathically: A question of heart, not a question of skin. *Applied Psychophysiology and Biofeedback*, *36*(3), 201–207. doi: 10.1007/s10484-011-9161-2

Oppenheim, A. V., Willsky, A. S., & Nawab, S. H. (1983). *Signals and systems (Vol. 2).* Englewood Cliffs, NJ: Prentice Hall.

Orini, M., Bailón, R., Mainardi, L. T., Laguna, P., & Flandrin, P. (2012). Characterization of dynamic interactions between cardiovascular signals by time-frequency coherence. *IEEE transactions on bio-medical engineering*, *59*(3), 663–673. doi: 10.1109/TBME.2011.2171959

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* New York, NY: Cambridge University Press.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? *Psychological Review*, *97*(3), 315–331. doi: 10.1037//0033-295X.97.3.315

Overbeek, T. J. M., van Boxtel, A., & Westerink, J. H. D. M. (2012). Respiratory sinus arrhythmia responses to induced emotional states: Effects of RSA indices, emotion induction method, age, and sex. *Biological Psychology*, *91*(1), 128–141. doi: 10.1016/j.biopsycho.2012.05.011

Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, *32*(3), 230–236. doi: 10.1109/tbme.1985.325532

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *Behavioral and Brain Sciences*, *5*(03), 407. doi: 10.1017/S0140525X00012759

Panksepp, J. (2008). The affective brain and core consciousness: How does neural activity generate emotional feelings? In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 47–67). New York: Guilford Press.

Panksepp, J., & Watt, D. (2011). What is basic about basic emotions? Lasting lessons from affective neuroscience. *Emotion Review*, *3*(4), 387–396. doi: 10.1177/1754073911410741

Peláez-Coca, M. D., Orini, M., Lázaro, J., Bailón, R., & Gil, E. (2013). Cross time-frequency analysis for combining information of several sources: Application to estimation of spontaneous respiratory rate from photoplethysmography. *Computational and Mathematical Methods in Medicine*(6), 1–8. doi: 10.1155/2013/631978

Picard, R. W. (1995). *Affective Computing.* (Tech. Rep. No. 321). MIT.

Picard, R. W. (1997). *Affective Computing.* Cambridge, MA: MIT Press.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(10), 1175–1191. doi: 10.1109/34.954607

Poláčková Šolcová, I., & Lačev, A. (2017). Differences in male and female subjective experience and physiological reactions to emotional stimuli. *International Journal of Psychophysiology*, *117*, 75–82. doi: 10.1016/j.ijpsycho.2017.04.009

Quigley, K. S., & Barrett, L. F. (2014). Is there consistency and specificity of autonomic changes during emotional episodes? Guidance from the Conceptual Act Theory and psychophysiology. *Biological Psychology*, *98*, 82–94. doi: 10.1016/j.biopsycho.2013.12.013

R Core Team. (2012). *R: A language and environment for statistical computing.* Computer software manual. Vienna, Austria. Retrieved from `http://www.R-project.org`

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology*, *61*(1), 5–18. doi: 10.1016/j.ijpsycho.2005.10.024

Rani, P., Liu, C., Sarkar, N., & Vanman, E. (2006). An empirical study of machine learning techniques for affect recognition in human–robot interaction. *Pattern Analysis and Applications*, *9*(1), 58–69. doi: 10.1007/s10044-006-0025-y

Rani, P., Sarkar, N., & Adams, J. (2007). Anxiety-based affective communication for implicit human-machine interaction. *Advanced Engineering Informatics*, *21*(3), 323–334. doi: 10.1109/TCAD.2006.11.009

Reed, R. G., Randall, A. K., Post, J. H., & Butler, E. A. (2013). Partner influence and in-phase versus anti-phase physiological linkage in romantic couples. *International Journal of Psychophysiology*, *88*(3), 309–316. doi: 10.1016/j.ijpsycho.2012.08.009

Reisenzein, R. (1994). Pleasure-arousal theory and the intensity of emotions. *Journal of Personality and Social Psychology*, *67*(3), 525–539. doi: 10.1037// 0022-3514.67.3.525

Reisenzein, R. (2000). Exploring the strength of association between the components of emotion syndromes: The case of surprise. *Cognition & Emotion*, *14*(1), 1–38. doi: 10.1080/026999300378978

Ritz, T., Thons, M., Fahrenkrug, S., & Dahme, B. (2005). Airways, respiration, and respiratory sinus arrhythmia during picture viewing. *Psychophysiology*, *42*, 568–578. doi: 10.1111/j.1469-8986.2005.00312.x

Rohrmann, S., & Hopp, H. (2008). Cardiovascular indicators of disgust. *International Journal of Psychophysiology*, *68*(3), 201–208. doi: 10.1016/ j.ijpsycho.2008.01.011

Rohrmann, S., Hopp, H., Schienle, A., & Hodapp, V. (2009). Emotion regulation, disgust sensitivity, and psychophysiological responses to a disgust-inducing film. *Anxiety Stress and Coping*, *22*(2), 215–236. doi: 10.1080/ 10615800802016591

Roseman, I. J., Spindel, M. S., & Jose, P. E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, *59*(5), 899–915. doi: 10.1037/0022-3514.59.5.899

Rosenberg, E. L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition & Emotion*, *8*(3), 201–229. doi: 10.1080/02699939408408938

Roth, W. T., Telch, M. J., Barr Taylor, C., & Stewart Agras, W. (1988). Autonomic changes after treatment of agoraphobia with panic attacks. *Psychiatry Research*, *24*(1), 95–107. doi: 10.1016/0165-1781(88)90145-X

Rottenberg, J., Ray, R., & Gross, J. J. (2007). Emotion elicitation using films. In J. A. Coan & J. J. B. Allen (Eds.), *Handbook of emotion elicitation and assessment* (pp. 9–28). New York: Oxford University Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. doi: 10.1037/h0077714

Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, *110*(3), 426–450. doi: 10.1037/0033-2909.110.3.426

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

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819. doi: 10.1037/0022-3514.76.5.805

Saarimäki, H., Gotsopoulos, A., Jääskeläinen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., ... Nummenmaa, L. (2016). Discrete neural signatures of basic emotions. *Cerebral Cortex*, *26*(6), 2563–2573. doi: 10.1093/cercor/bhv086

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379–399. doi: 10.1037/h0046234

Schäck, T., Muma, M., Feng, M., Guan, C., & Zoubir, A. (2017). Robust nonlinear causality analysis of non-stationary multivariate physiological time series. *IEEE Transactions on Bio-medical Engineering*, *99*, 1. doi: 10.1109/TBME.2017.2708609

Schäck, T., Sledz, C., Muma, M., & Zoubir, A. M. (2015). A new method for heart rate monitoring during physical exercise using photoplethysmographic signals. In *European Signal Processing Conference 2015.* Retrieved from `http://tubiblio.ulb.tu-darmstadt.de/71813/`

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition & Emotion*, *24*(7), 1153–1172. doi: 10.1080/02699930903274322

Schaefer, H. S., Larson, C. L., Davidson, R. J., & Coan, J. A. (2014). Brain, body, and cognition: Neural, physiological and self-report correlates of phobic and normative fear. *Biological Psychology*, *98*, 59–69. doi: 10.1016/j.biopsycho.2013.12.011

Scharf, L. L. (1991). *Statistical signal processing* (Vol.98 ed.). Reading, MA: Addison-Wesley.

Scherer, K. R. (1987). Toward a dynamic theory of emotion: The component process model of affective states. *Geneva Studies in Emotion and Communication*, *1*, 1–98.

Scherer, K. R. (1993). Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition & Emotion*, *7*(3-4), 325–355. doi: 10.1080/02699939308409192

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, *23*(7), 1307–1351. doi: 10.1080/02699930902928969

Scherer, K. R., Banziger, T., & Roesch, E. (Eds.). (2010). *Blueprint for affective computing: A sourcebook and manual.* Oxford: Oxford University Press.

Scherer, K. R., Dan, E., & Flykt, A. (2006). What determines a feeling's position in affective space? A case for appraisal. *Cognition & Emotion*, *20*(1), 92–113. doi: 10.1080/02699930500305016

Scherer, K. R., & Ekman, P. (Eds.). (1984). *Approaches to emotion.* Hillsdale, NJ: Lawrence Erlbaum.

Schienle, A., Stark, R., & Vaitl, D. (2001). Evaluative conditioning: A possible explanation for the acquisition of disgust responses? *Learning and Motivation*, *32*(1), 65–83. doi: 10.1006/lmot.2000.1067

Schienle, A., Walter, B., Stark, R., & Vaitl, D. (2002). Ein Fragebogen zur Erfassung der Ekelempfindlichkeit (FEE). *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis*, *31*(2), 110–120. doi: 10.1026//1616-3443.31.2.110

Schijvenaars, B. J., van Herpen, G., & Kors, J. A. (2008). Intraindividual variability in electrocardiograms. *Journal of Electrocardiology*, *41*(3), 190–196. doi: 10.1016/j.jelectrocard.2008.01.012

Shen, T. W., Tompkins, W. J., & Hu, Y. H. (2002). One-lead ECG for identity verification. In J. W. Clark (Ed.), *Proceedings of Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference* (pp. 62–63). doi: 10.1109/IEMBS.2002.1134388

Shumway, R. H., & Stoffer, D. S. (2006). *Time Series Analysis and Its Applications:*

With R Examples. New York, NY: Springer.

Sinha, R., & Parsons, O. A. (1996). Multivariate response patterning of fear and anger. *Cognition & Emotion*, *10*(2), 173–198. doi: 10.1080/026999396380321

Stark, R., Walter, B., Schienle, A., & Vaitl, D. (2005). Psychophysiological correlates of disgust and disgust sensitivity. *Journal of Psychophysiology*, *19*(1), 50–60. doi: 10.1027/0269-8803.19.1.50

Stemmler, G. (1989). Psychophysiological assessment: Conceptual, psychometric, and statistical issues. In G. Turpin (Ed.), *Handbook of clinical psychophysiology* (pp. 71–104). Chichester, New York: Wiley.

Stemmler, G. (1992). *Differential psychophysiology: Persons in situations.* Berlin, Germany: Springer-Verlag.

Stephens, C. L., Christie, I. C., & Friedman, B. H. (2010). Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biological Psychology*, *84*(3), 463–473. doi: 10.1016/j.biopsycho.2010.03.014

Stoica, P., & Moses, R. (2005). *Spectral analysis of signals.* Upper Saddle River, NJ: Pearson Education.

Strasser, F., Muma, M., & Zoubir, A. (2012). Motion artifact removal in ECG signals using multi-resolution thresholding. In *Proceedings of the 20th European Signal Processing Conference* (pp. 899–903). Piscataway, N.J: IEEE.

Swangnetr, M., & Kaber, D. B. (2013). Emotional state classification in patient–robot interaction using wavelet analysis and statistics-based feature selection. *IEEE Transactions on Human-Machine Systems*, *43*(1), 63–75. doi: 10.1109/TSMCA.2012.2210408

Sze, J. A., Gyurak, A., Yuan, J. W., & Levenson, R. W. (2010). Coherence between emotional experience and physiology: Does body awareness training have an impact? *Emotion*, *10*(6), 803–814. doi: 10.1037/a0020146

Tassinary, L. G., Cacioppo, J. T., & Vanman, E. J. (2007). The skeletomotor system: Surface electromyography. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 163–199). Cambridge: Cambridge University Press.

Thompson, M., & Thompson L. (2003). *The Biofeedback Book: An Introduction to Basic Concepts in Applied Psychophysiology.* Wheat Ridge, CO: Association for Applied Psychophysiology and Biofeedback.

Tomkins, S. (1962). *Affect, imagery, consciousms: Vol. 1, The positive affects.*

New York: Springer.

Touroutoglou, A., Lindquist, K. A., Dickerson, B. C., & Barrett, L. F. (2015). Intrinsic connectivity in the human brain does not reveal networks for 'basic' emotions. *Social Cognitive and Affective Neuroscience*, *10*(9), 1257–1265. doi: 10.1093/scan/nsv013

Tracy, J. L., & Randles, D. (2011). Four models of basic emotions: A review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review*, *3*(4), 397–405.

Tukey, J. W. (1979). Robust techniques for the user. *Robustness in Statistics*, *103*, 106.

Valstar, M. F., & Pantic, M. (2012). Fully automatic recognition of the temporal phases of facial actions. *IEEE Trans. Syst., Man, Cybern. B (IEEE Transactions on Systems, Man, and Cybernetics, Part B)*, *42*(1), 28–43. doi: 10.1109/TSMCB.2011.2163710

Vapnik, V. N. (1998). *Statistical learning theory.* New York: Wiley.

Vaseghi, S. V. (2008). *Advanced digital signal processing and noise reduction.* New York, NY: Wiley.

Venkatesh, N., & Srinivasan, J. (2010). Human electrocardiogram for biometrics using DTW and FLDA. In *Proceedings of the 20th International Conference on Pattern Recognition* (pp. 3838–3841). doi: 10.1109/ICPR.2010.935

Verduyn, P., van Mechelen, I., Tuerlinckx, F., Meers, K., & van Coillie, H. (2009). Intensity profiles of emotional experience over time. *Cognition & Emotion*, *23*(7), 1427–1443. doi: 10.1080/02699930902949031

Wagner, J., Jonghwa, K., & Andre, E. (2005). From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. In *Proceedings of the 27th Annual International Conference of the Engineering Medicine and Biology Society* (pp. 940–943). doi: 10.1109/ICME.2005.1521579

Walecki, R., Rudovic, O., Pavlovic, V., & Pantic, M. (2017). Variable-state Latent Conditional Random Field models for facial expression analysis. *Image and Vision Computing*, *58*, 25–37. doi: 10.1016/j.imavis.2016.04.009

Walter, S., Kim, J., Hrabal, D., Crawcour, S. C., Kessler, H., & Traue, H. C. (2013). Transsituational individual-specific biopsychological classification of emotions. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *43*(4), 988–995. doi: 10.1109/TSMCA.2012.2216869

Wang, J., Neskovic, P., & Cooper, L. N. (2007). Improving nearest neighbor rule with a simple adaptive distance measure. *Pattern Recognition Letters*, *28*(2), 207–213. doi: 10.1016/j.patrec.2006.07.002

Waugh, C. E., & Schirillo, J. A. (2012). Timing: A missing key ingredient in typical fMRI studies of emotion. *Behavioral and Brain Sciences*, *35*(03), 170–171. doi: 10.1017/S0140525X11001646

Wei, W., & Jia, Q. (2016). Weighted feature gaussian kernel SVM for emotion recognition. *Computational Intelligence and Neuroscience*, 1–7. doi: 10.1155/2016/7696035

Wenger, M. A. (1942). A study of physiological factors: the autonomic nervous system and the skeletal musculature. *Human Biology*, *14*, 69–84.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience*, *18*(1), 411–418.

White, L. B., & Boashash, B. (1990). Cross spectral analysis of nonstationary processes. *IEEE Transactions on Information Theory*, *36*(4), 830–835. doi: 10.1109/18.53742

Wiem, M. B. H., & Lachiri, Z. (2017). Emotion classification in arousal valence model using MAHNOB-HCI database. *International Journal of Advanced Computer Science and Applications*, *8*(3), 318–323.

Wilhelm, F. H., & Grossman, P. (2010). Emotions beyond the laboratory: Theoretical fundaments, study design, and analytic strategies for advanced ambulatory assessment. *Biological Psychology*, *84*(3), 552–569. doi: 10.1016/j.biopsycho.2010.01.017

Wöllmer, M., Kaiser, M., Eyben, F., Schuller, B., & Rigoll, G. (2013). LSTM-Modeling of continuous emotions in an audiovisual affect recognition framework. *Image and Vision Computing*, *31*(2), 153–163. doi: 10.1016/j.imavis.2012.03.001

Wu, L., Winkler, M. H., Andreatta, M., Hajcak, G., & Pauli, P. (2012). Appraisal frames of pleasant and unpleasant pictures alter emotional responses as reflected in self-report and facial electromyographic activity. *International Journal of Psychophysiology*, *85*(2), 224–229. doi: 10.1016/j.ijpsycho.2012.04.010

Wu, T. F., Lin, C. J., & Weng, R. C. (2004). Probability estimates for multi-class classification by pairwise coupling. *Journal of Machine Learning Research*, *5*, 975–1005.

Yasuma, F. (2004). Respiratory sinus arrhythmia - Why does the heartbeat synchronize with respiratory rhythm? *CHEST Journal*, *125*(2), 683. doi: 10.1378/chest.125.2.683

Ye, C., Coimbra, M. T., & Kumar, B. V. K. V. (2010). Investigation of human identification using two-lead Electrocardiogram (ECG) signals. In *Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems* (pp. 1–8). doi: 10.1109/BTAS.2010.5634478

Zhou, F., Qu, X., Jiao, J., & Helander, M. G. (2014). Emotion prediction from physiological signals: A comparison study between visual and auditory elicitors. *Interacting with Computers*, *26*(3), 285–302. doi: 10.1093/iwc/iwt039

Zong, C., & Chetouani, M. (2009). Hilbert-Huang transform based physiological signals analysis for emotion recognition. In *Proceedings of the 9th IEEE International Symposium on Signal Processing and Information Technology* (pp. 334–339). doi: 10.1109/ISSPIT.2009.5407547

Zoubir, A. M., Koivunen, V., Chakhchoukh, Y., & Muma, M. (2012). Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts. *IEEE Signal Processing Magazine*, *29*(4), 61–80. doi: 10.1109/msp.2012.2183773

Zucker, T. L., Samuelson, K. W., Muench, F., Greenberg, M. A., & Gevirtz, R. N. (2009). The effects of respiratory sinus arrhythmia biofeedback on heart rate variability and posttraumatic stress disorder symptoms: A pilot study. *Applied Psychophysiology and Biofeedback*, *34*(2), 135–143. doi: 10.1007/s10484-009-9085-2

List of Manuscripts

Manuscript A:
Kelava, A., Muma, M., Deja, M., Dagdagan, J. Y., & Zoubir, A. M. (2015). A new approach for the quantification of synchrony of multivariate non-stationary psychophysiological variables during emotion eliciting stimuli. *Frontiers in Psychology, 5.* DOI: 10.3389/fpsyg.2014.01507.

Die Verfasserin wirkte als Koautorin bei der Konzeption und der Erstellung des Artikels tatkräftig mit. Insbesondere der theoretische Hintergrund sowie die Diskussion wurden von der Verfasserin geschrieben, die anderen Teile durch Korrekturlesen mit gestaltet. Die Verfasserin war beim Datenerhebungs- und auswertungsprozess beteiligt. Der Artikel wurde in der Zeitschrift Frontiers in Psychology veröffentlicht.

Manuscript B:
Deja, M., Muma, M., Hoppe, D. & Kelava, A. (to be revised and resubmit). Synchrony of psychophysiological parameters in disgust. *International Journal of Psychophysiology*, 56 pages.

Die Verfasserin war als Erstautorin maßgeblich für die Konzeption sowie das Abfassen des oben genannten Artikels verantwortlich. Sie war aktiv in jeder Phase der Datenerhebung eingebunden (Versuchsaufbau, Auswahl des Stimulus Materials und der Fragebögen, Ablauf des Experimentes, Rekrutierung der Probanden, Datenerhebung im Labor). Ferner war sie für das Zusammentragen der Daten verantwortlich und in Kooperation mit dem Fachbereich Elektro-Technik der TU Darmstadt auch an der Auswertung der physiologischen und subjektiven Daten maßgeblich beteiligt. Die Erstautorin hat die Einführung, die Ergebnisse und die Diskussion sowie Abschnitte der Methoden federführend verfasst, das Erstellen und Bearbeitung von Abbildungen und Tabellen mit eingeschlossen. Der Artikel wurde in einem Peer Review Verfahren der Zeitschrift International Journal of Psychophysiology aufgenommen.

Manuscript C:
Deja, M., Muma, M., Richter, S., Entringer, T., Hoppe, D. & Kelava, A. (submitted). Person- and emotion classification of fear using peripheral physiological data. *Psychological Test and Assessment Modeling*, 46 pages.

Die Verfasserin war als Erstautorin maßgeblich für die Konzeption sowie das Abfassen des oben genannten Artikels verantwortlich. Sie war aktiv in jeder Phase der Konzeption und Datenerhebung (Versuchsaufbau, Auswahl des Stimulus Materials und der Fragebögen, Ablauf des Experimentes, Rekrutierung der Probanden, Datenerhebung im Labor) eingebunden. Die Auswertung wurde erneut in Kooperation mit dem Fachbereich Elektro-Technik der TU Darmstadt vorgenommen. Die Einleitung, Ergebnisse und Diskussion sowie die meisten Abschnitte der Methode wurden federführend von der Erstautorin verfasst. Der Artikel wurde in der Zeitschrift Psychological Methods and Assessment eingereicht.

Scientific CV

Marlene Dejá (geb. Schmidt), Dipl.-Psych.
Date of Birth: 04/09/1985
Place of Birth: Frankfurt am Main, Germany
Nationality: German

09/2011-04/2016	Research Assistant, Methods Division, Technische Universität Darmstadt
07/2011	Diploma in Psychology, Goethe-Universität Frankfurt
10/2007-07/2011	Student Assistant, Methods Division, Division of Clinical Psychology, Goethe-Universität Frankfurt
04/2006-07/2011	Studies of Psychology, Goethe-Universität Frankfurt
06/2005	Diploma, Schule am Ried, Frankfurt

Obligatory Declaration

I declare that I have developed and written the enclosed doctoral thesis entitled "Response Synchrony and Response Patterning of Psychophysiological Parameters in Emotion" completely by myself, and have not used sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked. This thesis was not used in the same or in a similar version to achieve an academic grading or is being published elsewhere.

Datum, Ort   Marlene Dejá, Dipl.-Psych.

A. Appendix

Table A1.

Five Most Disgusting Pictures Rated via the Rating Dial in the Prestudy

		Rating dial	
Picture	Content	M	SE
IAPS no. 3150	A half ripped-off finger	4.93	0.80
IAPS no. 9301	A toilet with excrement	4.59	0.65
IAPS no. 3100	A face with a burn	4.35	0.81
Diarrhea	A toilet with excrement	4.01	0.69
Vomit	A picture of puke in a car	3.57	0.56

Note. M = mean; SE = standard error.

Table A2.

Manipulation Check: Paired t Tests for the Mean of the Mean Difference Scores

DAS Scale	M_{M_d}	SE	t	p	Cohen's d
disgust-fear	1.78-0.37 = 1.42	.13	10.76	.00**	1.66
disgust-anger	1.78-0.17 = 1.61	.16	9.98	.00**	1.54
disgust-sadness	1.78-0.23 = 1.55	.16	9.46	.00**	1.46
disgust-surprise	1.78-0.22 = 1.56	.17	9.13	.00**	1.41

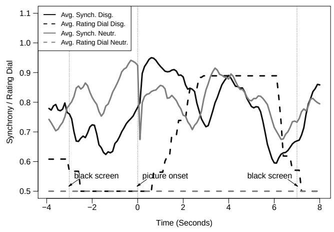
Note. DAS = Differential Affect Scale. M_{M_d}= mean of the means of differences (disgust $t_2 - t_1$ - emotion $t_2 - t_1$), SE = standard error, p-values = p-values corrected by the *Bonferroni* correction. N = 42; df = 41. **p < .01.

Table A3.

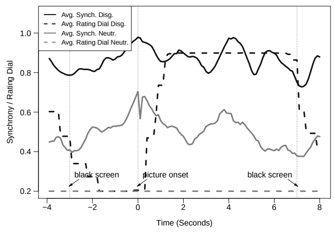
Means of the Differences, Standard Errors, t Values, p-values, and Cohen's d of the Paired t Tests for the Physiological Parameters: Heart Rate (HR), Heart Rate Variability (HRV), Skin Conductance Response (SCR), and Inspiratory Depth (ID)

	M_D	SE	t	p	Cohen's d
HR	-6.08	0.79	-7.73	.00**	-1.19
HRV	0.01	0.00	8.57	.00**	1.32
SCR	0.11	0.04	2.61	.04*	0.40
ID	-0.27	0.20	-1.35	.74	-0.21

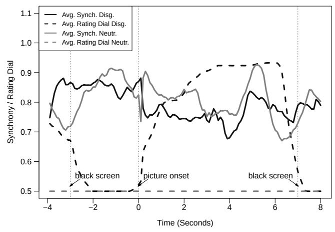
Note. M_D = Mean of the differences (disgust t_2 - neutral t_1); SE= standard error of the mean of the differences; p-values= p-values corrected by the *Bonferroni* correction. N = 42; *df* = 41. *p < .05. **p < .01.



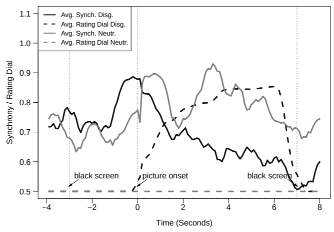
(a) Participant 1



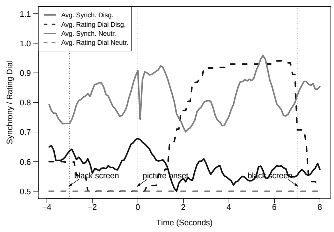
(b) Participant 2



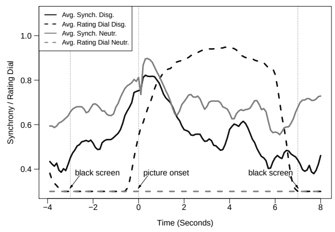
(c) Participant 3



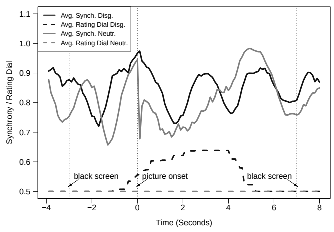
(d) Participant 4



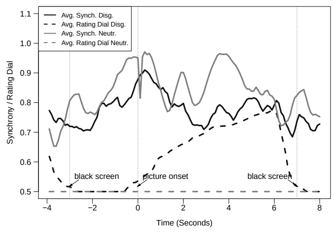
(e) Participant 5



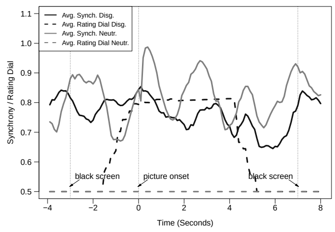
(f) Participant 6



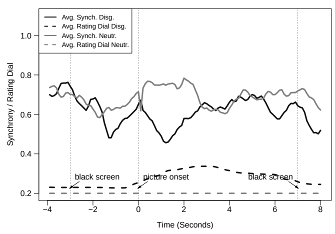
(g) Participant 7



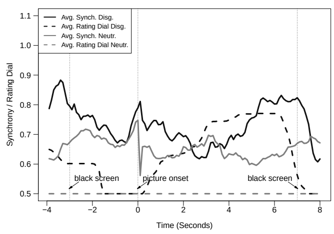
(h) Participant 8



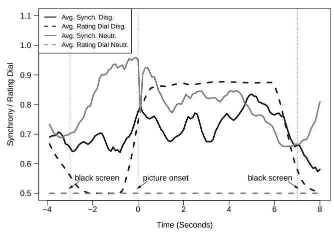
(i) Participant 9



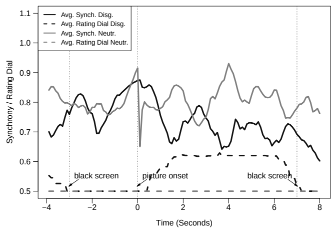
(j) Participant 10



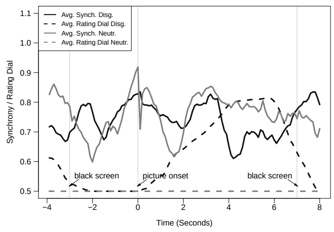
(k) Participant 11



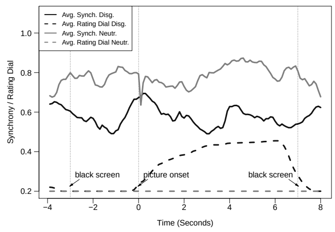
(l) Participant 12



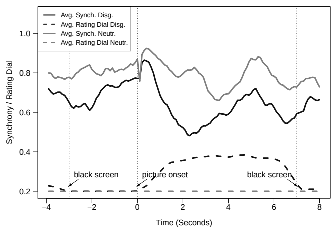
(m) Participant 13



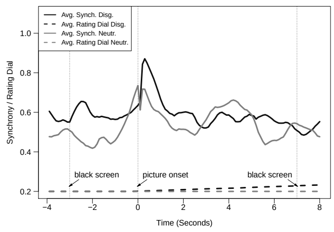
(n) Participant 14



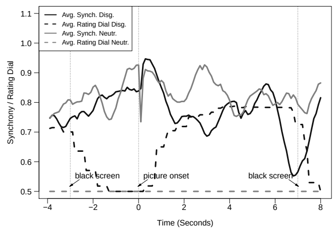
(o) Participant 15



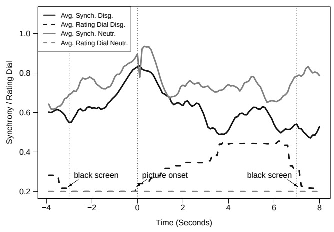
(p) Participant 16



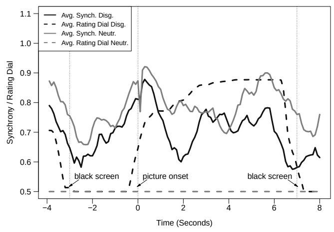
(q) Participant 17



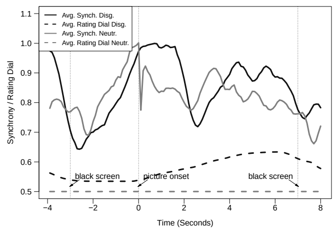
(r) Participant 18



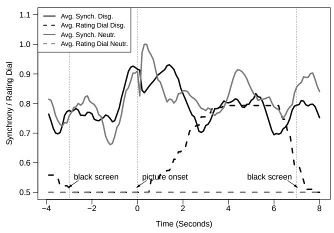
(s) Participant 19



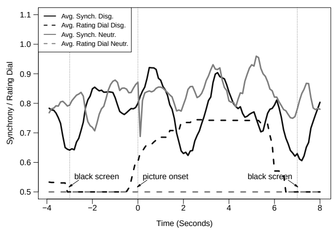
(t) Participant 20



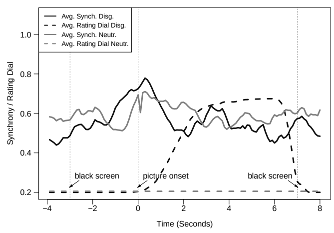
(u) Participant 21



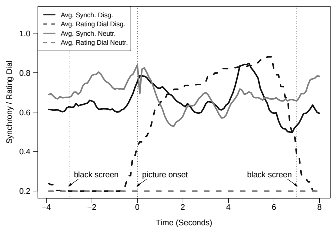
(v) Participant 22



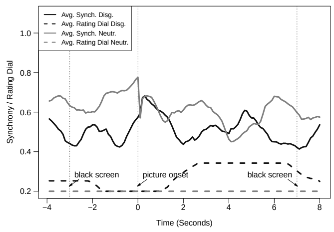
(w) Participant 23



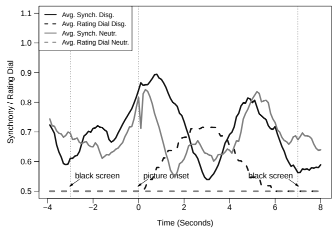
(x) Participant 24



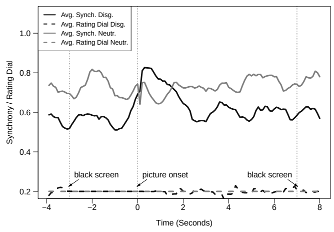
(y) Participant 25



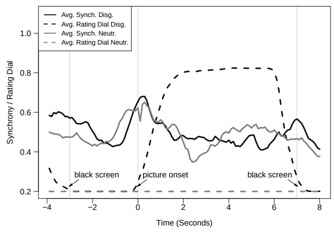
(z) Participant 26



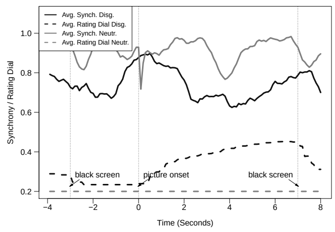
(aa) Participant 27



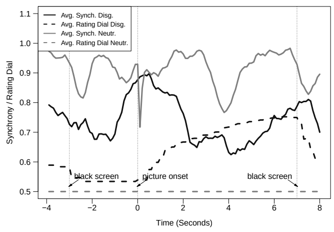
(ab) Participant 28



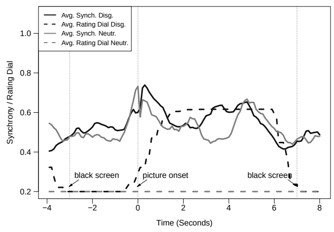
(ac) Participant 29



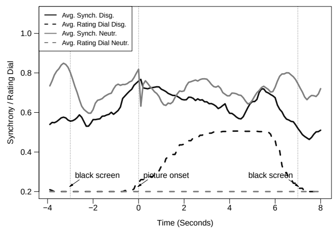
(ad) Participant 30



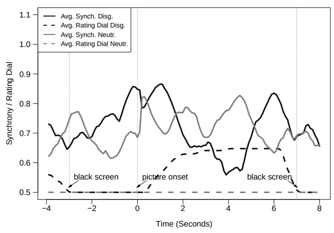
(ae) Participant 31



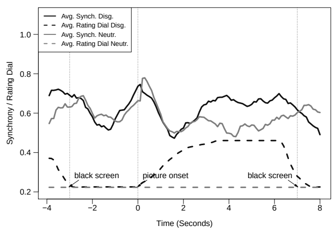
(af) Participant 32



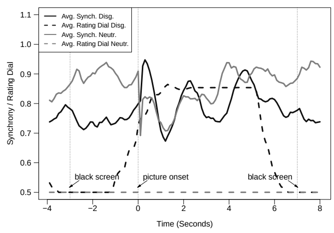
(ag) Participant 33



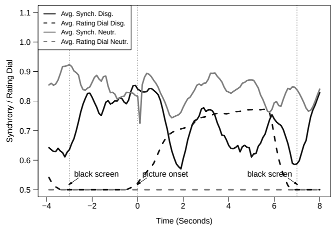
(ah) Participant 34



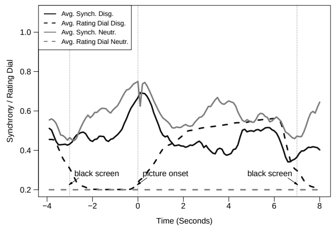
(ai) Participant 35



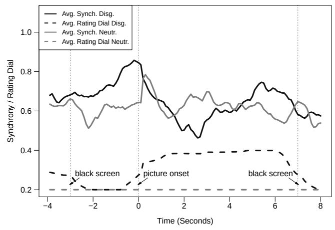
(aj) Participant 36



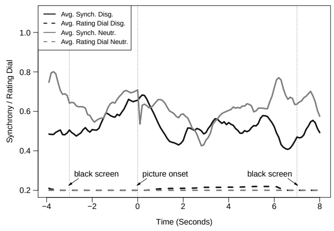
(ak) Participant 37



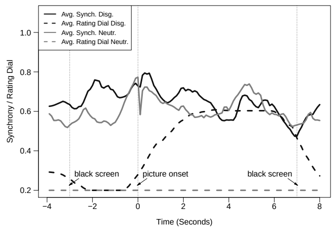
(al) Participant 38



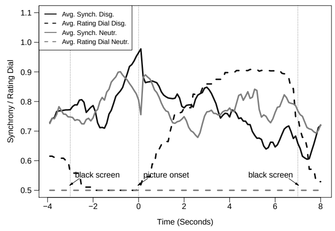
(am) Participant 39



(an) Participant 40

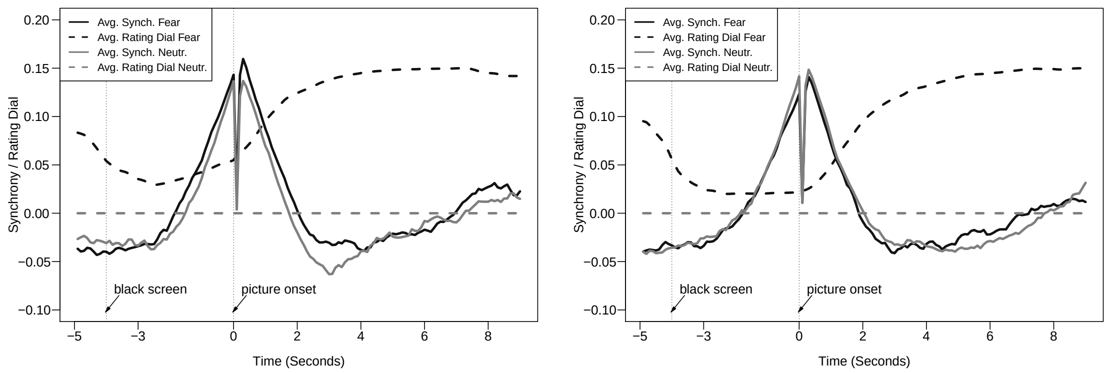


(ao) Participant 41



(ap) Participant 42

Figure A1. The average intraindividual time course of the latent physiological synchrony (Synch.) and average continuous rating dial values while viewing disgusting (Disg.) and neutral (Neutr.) picture sequences for all participants. The y-axis contains the latent physiological synchrony and the continuous rating dial values. The x-axis represents the time in seconds. The point in time t = 0 s corresponds to the moment when a new picture (either neutral or disgusting) was shown to the participant. This picture was shown for 7 s. Before and after the picture, a black screen appeared for a duration of 3 s (i.e., a black screen appeared at t = -3 s and at t = 7 s), t < -3 s and t > 7 s correspond to the previous and subsequent pictures, respectively.



(a) Measurement point 1 (b) Measurement point 2

Figure A2. Average (Avg.) time course of the overall physiological synchrony (Synch.) for neutral (Neutr.) and fearful pictures at two measurement points with a time interval of 6 weeks between them. Synchrony was averaged across all individuals for the neutral and fearful pictures separately. The same type of averaging across individuals was performed for the rating dial values, resulting in two average time courses. The absolute level of synchrony was subtracted by centering the data. The x-axis represents the time in seconds. The point in time $t = 0$ s corresponds to the moment when a new picture (either neutral or fearful) was shown to the participant. This picture was shown for 15 s, but only the first 8 s of the picture are displayed. Before and after the picture, a black screen appeared for a duration of 4 s (i.e., a black screen appeared at $t =$ -4 s and at $t = 15$ s).