UNIVERSITÀ DEGLI STUDI
DI TRENTO

DEPARTMENT OF INFORMATION ENGINEERING AND COMPUTER SCIENCE
ICT International Doctoral School

IMPLICIT HUMAN-COMPUTER INTERACTION: TWO COMPLEMENTARY APPROACHES

Julia Wache

Advisor:

Prof. Nicu Sebe

Università degli Studi di Trento

April 2016

Abstract

One of the main goals of Human Computer Interaction (HCI) is to improve the interface between users and computers: interfacing should be effortless and easy to learn. In this thesis, we pursue this goal, aiming to reduce the stress of users and increase their wellbeing. We work on two different but complementary approaches: (i) Automatic assessment of users inner psychological state, so as to enhance computer-human interaction; and (ii) Information presentation in a comprehensive manner, with no stress added by devices and applications when delivering information. Not only computers should understand their users, but also users should easily understand the information given by computers. For the first approach we collected physiological and psychological data from people exposed to emotional stimuli. We created a database, and made it freely available to the community, for further use in research on automated detection of the differences in the inner states of users. We employed the data for predicting both the emotional state of users and their personality traits. For the second approach, we investigated two devices that intend to provide comprehensible feedback easily. First we discuss how to utilize a breathing sensor that informs its users on their current physiological state and on how to decrease the stress in daily life by adapting their breathing patterns. Here we investigated general criteria on how to develop systems that are easily understandable. The second device was a tactile belt. We analyze the belt as a solution that provides comprehensive guidance information in navigation contexts, and that does not require cognitive effort. The belt uses localized tactile stimulation to transmit directional information. By employing the tactile sense it can augment or even replace the information normally received through eyes and ears. Finally, we discuss opportunities for future applications of our research, and conclude with a summary of our contributions to HCI: transmitting information from humans to machines and vice versa.

Keywords
personality, affective computing, emotion recognition, physiological signals, tactile signals, navigation

Acknowledgements

When I was thinking about a PhD position four years ago, I was not immediately considering Computer Science as my background is in Cognitive Science. The direction my PhD would take was not very clear in the first year. Luckily, I found my advisor Prof. Nicu Sebe. I am very grateful for his guidance and constant support during my adventurous PhD journey. Thanks for introducing me to my dear colleague Mojtaba Khomami Abadi, who taught me most of the technical knowledge I acquired and was always ready to help and support me. Also thanks to my other colleagues that worked with me, especially Ramanathan Subramanian.

I thank the European Institute of Innovation & Technology that I could be part of the EIT programme where I learned a lot of additional entrepreneurial skills that researchers often don't get the chance to learn. Transferring research results into useful applications for daily life became my goal for both my PhD and my future life, at the latest after successfully participating in the student competition in Virginia Tech.

I thank Neema Moraveji from Spire, for inviting me to San Francisco and giving me the exciting inside into a startup.

I am especially happy that I got the opportunity to do a further internship at the startup feelSpace. Special thanks for the great collaboration to Silke, Jessika and Susan who welcomed me to the feelSpace family and wait for me to come back and join their venture.

Trento would not have been the great experience it was without the people that made me feel home. First of all, I want to thank Marcos, who always was there for me and made these years special. Also a great impact had my flatmates, the Costanzi family, with the current, former and extended members, Marco, Francisco, Marijo, Cristhian and of course Ruzanna. They were always *ready to listen to my PhD-related problems, counsel me and provide the necessary distractions. The Paraguayan community here in Trento made me feel as one of them (and they are the reason I learned more Spanish than Italian here). I also thank all my other friends who supported me throughout the journey.*

Last but not least I want to thank my family. Thanks for the loving support of my parents who taught me to see the world with amazement and planted the seed of the curiosity that keeps me wanting to learn more about the wonders of life and nature. Finally, am very grateful to have my sister Susan who always knows the right thing to say and do to me feel good.

Contents

Chapter 1

Introduction

In the past decades new technologies became a constant and important part of human life being used in many areas of interest. The typical life and work style of people have changed resulting in people spending more time interacting with computers and other smart devices. To facilitate the experience and improve the interaction with users, Human Computer Interaction (HCI) research aims constantly at improving the interaction processes with new technologies.

Unlike those of computers, not all actions of humans are predictable or follow logical rules; in fact one important factor of human behavior is emotion. External cues from the environment as well as internal states induce emotions that can have a substantial influence on the peoples experience and subsequent actions. The information a computer system has on the assumed emotional status of a user is mostly limited. Therefore, it is useful for a system to learn from interactions with users and their behavior about their inner state and current emotions. If the system is able to detect and understand human emotions without forcing the user to explicitly name his or her emotions (e.g. by providing written feedback), the system can adapt then the consequent interactions accordingly.

This chapter presents the outline of the approach and methods of the research conducted throughout this thesis.

1.1 Research approaches

To enhance interactions between users and their computers two different aspects are important: (1) enable computers to assess user's differences automatically to improve the understanding of their current inner state and (2) make technology provide information in an easy accessible way that avoids stress and cognitive overload of users.

1.1.1 Technology understanding humans

One important factor influencing human behavior is emotion. Emotions are psychological phenomena but they also have an impact on physiological signals that can be measured and analyzed (e.g. [142]). Physiological changes are induced by the Limbic System, the part of the brain controlling the autonomic nervous system. The Limbic System is involved in emotional experience and sensory input processing [72]. Hence, the current emotional state may influence the perception of a situation and the corresponding behavioral actions of people [134, 172]. Also further internal factors, such as personality, can have an impact. Therefore one challenge for contemporary HCI systems is the ability to understand the influence that factors like the personality traits or the environment have on the emotional state of a user and to use this information to predict consequential behavioral actions. For example, in an e-learning scenario the computer might adapt the shown content depending on how easy (boring) or difficult (stressing) it is perceived by the user.

To investigate emotional states and what influences them, emotional videos are preferably used because multimedia contents such as movies intend to induce certain emotions in viewers to create a specific experience [176]. One way to understand the current emotional state of users while watching emotional stimuli is to interrupt the users and directly ask them for an evaluation. Newer methods avoid interrupting the user. Instead they measure users' implicit reactions in emotional situations and extract the affective state from the collected data. One method is to record physiological responses and facial expressions during exposure to emotional content [149, 89]. In this way we intend to solve the first problem - automated emotion and personality recognition.

Successful applications could cover the area of computer games or learning platforms where the user experience can be enhanced significantly when the system decreases the stress level, hence, the negative emotions. These applications should avoid frustration and should increase their success by adapting the speed and the difficulty of the tasks according to the users performance.

1.1.2 Humans understanding technology

The other side of an effective communication between humans and their technological devices is the ability to provide information in an easily understandable way making the processing of data effortless for the users. In a world of increasing information overload, people are often overwhelmed by the huge amount of data to be processed. Most input is usually received over the visual and auditory channels. In daily life people continuously encounter an endless stream of sounds and images which can lead to cognitive overload resulting in less effective performance in current tasks [182]. To relieve the visual and auditory senses, a further sense that can be employed for information input is the tactile sense. The sense of touch is suitable to transmit simple information in a discrete way, e.g. a mobile phone that vibrates to transmit information to its owner without attracting attention from other people close by.

We analyze a tactile alternative to common output devices (visual displays, loudspeakers, earphones) to provide precise information during navigational tasks. As a direct application outside of the lab, tactile belts can, in combination with a smartphone, be used for navigation in unknown environments. Current navigation aids are mostly based on audio or visual instructions. Normally navigating in a new environment requires eyes and ears to be focused on the surroundings, e.g. in a urban area pedestrians need to attend to traffic and evade obstacles or other people. The usage of tactile signals elegantly circumnavigates an overload of the visual and auditory sensory modalities as it uses the tactile sense to provide information. This can enhance the effective navigation for people requiring special assistance such as visually impaired or elderly people, but it can also improve navigation experience for people engaged in activities that require their hands (cyclists) or their eyes and ears to be free for other tasks, for example athletes or tourists. In summary, tactile devices can be applied in real world applications to facilitate information reception and understanding.

1.2 Research methodologies

When interacting with intelligent systems an important aspect is to take the contextual situation of the user into account. Therefore, adapting the interaction possibilities between computer and user according to the users mental state can facilitate understanding between humans and machines.

In order to reach the goal of automated response to emotional states, the emotions need to be detected first. Emotions are psychological phenomena but they also have physiological concomitants that can be measured and analyzed (e.g. [148]). The challenge is to accurately detect physiological changes caused by affective responses and understand the dependencies between physiological signals, the emotions and the personality of users.

To analyze the physiological reactions we need to induce similar emotions in many people. This is exactly what multimedia contents such as movies intend to do. They evoke certain emotions in viewers to create an enjoyable experience. In fact, often the emotional experience stimulated is the aspect people value when watching a movie [9].

Important aspects of the mental state are current emotions and personality traits. Emotions are correlated to implicit responses [105] but it is yet to investigate in more depth how personality affects this relationship. Additionally, there is an increasing need of easy information transfer from machines to humans to avoid information overload. To achieve these goals we work on the following problems:

1. We need to automatically detect the mental state of users such as emotions and personality in order to improve interactions with intelligent systems. Therefore, we correlate physiological signals to emotions and personality traits and try predicting both, emotional states and personality traits, from those signals to understand how the goal of automated detection of inner states can be achieved best.

2. The second problem is to find a good way to communicate information to users and make it easily understandable. Information need to be transmitted in a complex context even when eyes and ears are already occupied by other tasks that require attention. We investigate whether tactile signals can overcome limitations of attention and be used to transmit information intuitively. We investigate how a tactile device in form of a belt can be best used to display information for navigation purposes and whether we can create an easy understanding of directional indications to reduce the required mental effort in orientation.

In the end we discuss how knowledge about the mental state and the personality of a user can be utilized by a system to provide information in a personalized way adapted to individual needs. We review how the results could be applied to future applications.

1.3 Outline of the thesis

Our general goal is to improve the intuitiveness of new technological developments to make the interaction with them easier and more effective. There are two complementary approaches that we address. To investigate the problems mentioned in the last section we developed and applied two main lines of investigation. In Figure 1.1 we show a schematic overview of the thesis. We highlight the two directions of communications in HCI with two different colors. For each flow of information we indicate the sender, the type of information provided, the supporting device (in grey), and the outcome that reaches the receiver. In more detail we did the following:

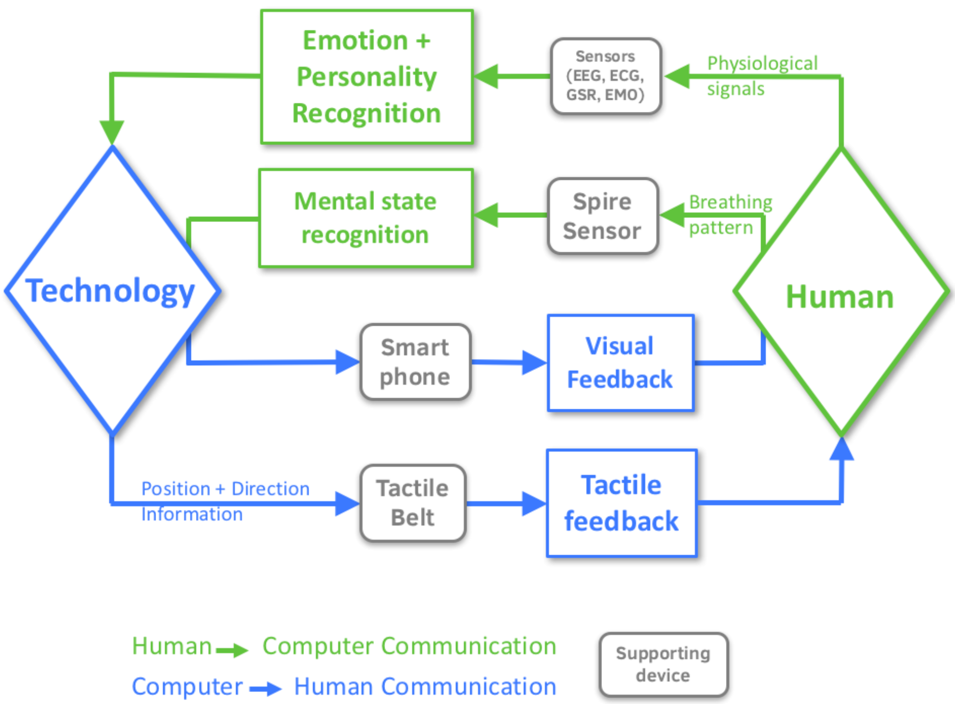


Figure 1.1: This schema provides an overview about this thesis. The first part focuses on the transfer of Information from Human to Computer, while the second part concentrates on the transfer of information from technology to humans.

In the next two chapters we aim at better understanding the user. Therefore, we investigate how to predict the inner mental state automatically, without users' having to actively provide information themselves, by measuring implicit signals. We then use those to predict both, the affective state and the personality traits.

We employ facial expressions and physiological signals, implicit signals, as they are mostly unconscious and difficult to influence. Therefore, we use sensors measuring their physiological changes while subjects watch multimedia content that elicits different emotions. With pattern recognition and machine learning approaches we then work on predicting emotions as well as personality traits from the collected measurements. We conduct two main experiments explained in chapter 2 and 3 where we collect physiological data, train classifiers to predict both, their affective reactions to emotional stimuli and personality traits, and use correlation methods to understand their connections. Both chapters are based on papers we published at respected conferences [1, 174]. From the study in chapter 2 we additionally prepared a dataset that is under review to be published for the community [173].

In the next project, we investigate how breathing can be used as an indicator for physiological states such as stress (chapter 4). We work on developing a sensor able to measure breath rates and an smartphone application to provide useful feedback. We evaluate how it can be used as calming technology to effectively reduce stress in users.

Finally we work on a use case investigating how to provide relevant information to the users in an easily comprehensible way. For this scenario we developed a tactile belt as a navigation device. It provides feedback via a tactile display to indicate directions. In chapter 5 we review the benefits of using tactile signals and the advantages of the tactile belt as navigation device, in particular for more challenged users such as the visually impaired. We investigate the form of tactile feedback necessary to optimize understanding of the tactile signal. We perform a study to validate the prototype of the tactile belt and detect differences in navigation performance for two different versions of displaying directional information. Further we investigate performance changes of subjects when under cognitive load.

In the last chapter (chapter 6) we discuss the results of our research and put them in the context of possible applications. We conclude by summarizing our contributions and pointing out the limitations of our research.

Chapter 2

ASCERTAIN: A Multimodal Affective Database for Personality Assessment

We present ASCERTAIN[1] *– a multimodal databa**AS**e for impli**C**it p**ER**sonali**T**y and **A**ffect recognit**I**o**N** using commercial physiological sensors. To our knowledge, ASCERTAIN is the first database to connect personality traits and emotional states via physiological responses. ASCERTAIN contains big-five personality scales and emotional self-ratings of 36 users along with their Electroencephalogram (EEG), Electrocardiogram (ECG), Galvanic Skin Response (GSR) and facial activity data, recorded using commercially available sensors while viewing affective movie clips. We first examine relationships between users' affective ratings and personality scales in the context of prior observations, and then study linear and non-linear physiological correlates of emotion and personality. Our analysis suggests that the emotion–personality relationship is better captured by non-linear rather than linear statistics. We finally attempt binary emotion and personality trait recognition using physiological features. Ex-*

[1]This chapter is mainly based on the following two papers: 1. Julia Wache, Ramanathan Subramanian, Mojtaba Khomami Abadi, Radu-Laurentiu Vieriu, Nicu Sebe, and Stefan Winkler. Implicit User- centric Personality Recognition Based on Physiological Responses to Emotional Videos. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pages 239-246, 2015 [174]. 2. Julia Wache, Ramanathan Subramanian, Mojtaba Khomami Abadi, Radu-Laurentiu Vieriu, Nicu Sebe, and Stefan Winkler. ASCERTAIN: A Multimodal Affective Database for Personality Assessment. Submitted to IEEE Transactions on Affective Computing (TAC) [173].

perimental results cumulatively confirm that personality differences are better revealed while comparing user responses to emotionally homogeneous videos, and above-chance recognition is achieved for both affective and personality dimensions.

2.1 Introduction

Despite rapid advances in Human-computer Interaction (HCI) and continuous effort to improve user experience with computer systems, the need for systems to *recognize* and *adapt* to the **affective state** of users has been widely acknowledged. While affect is an important component influencing human behavior, nevertheless it a highly subjective phenomenon that is influenced by contextual and psychological factors including **personality**.

The personality–affect relationship has been actively studied ever since a correlation between the two was proposed in Eysenck's personality model [52]. Eysenck supposed that Extraversion, the personality dimension that describes a person as either talkative or reserved, is accompanied by low cortical arousal, i.e., extraverts require more external stimulation than introverts. His model also proposed that neurotics, characterized by negative feelings such as anxiety, are more sensitive to external stimulation and become easily upset or nervous due to minor stressors.

Many affective studies have attempted to validate and extend Eyesenk's findings. Some have used explicit user feedback in the form of affective self-ratings [127, 85], while others have measured physiological signals to acquire implicit user responses such as Electroencephalogram (EEG) activity [154] and heart rate [44]. However, few works have investigated affective correlates of traits other than Extraversion and Neuroticism. Social psychology studies have examined personality mainly via non-verbal social behavioral cues (see [171]

Table 2.1: Summary of the ASCERTAIN database.

Number of Participants	36
Number of Videos	36
Video Length	51–128 seconds ($\mu \pm \sigma$ = 80 $\pm$ 20)
Self-reported ratings	Arousal, Valence, Engagement Liking, Familiarity
Personality Scales	Extraversion, Agreeableness Conscientiousness, Neuroticism, Openness
Physiological signals	ECG, GSR, Frontal EEG, Facial features

for a review), but few works have modeled personality traits based on emotional behavior.

To facilitate research on emotion recognition from physiological signals researchers need large datasets comprising physiological and emotional data from many subjects, but due to high cost of time, equipment and interdisciplinary knowledge necessary to compile such datasets, there is only a limited number available [149] and no currently published data set includes personality data.

We created the dataset ASCERTAIN[2] which contains measures of Electrocardiogram (ECG), Galvanic skin response (GSR), facial expressions and Electroencephalography (EEG) in response to 36 affective video stimuli. This work includes the first publicly available dataset that also includes personality scores. It builds on [174] and examines the influence of personality differences on users' affective states. Hence, we investigate whether affective physiological responses are indicative of users' personality traits. ASCERTAIN contains personality scores and emotional self-ratings of 36 users in addition to their affective physiological responses (overview of the dataset in Table 2.1).

We utilize ASCERTAIN to (i) understand the relation between *emotional attributes* and *personality traits*, and (ii) characterize both via users' physiological

[2]Available at: http://mhug.disi.unitn.it/index.php/datasets/ascertain/

responses.

We specifically designed a movie-based study as movie scenes effectively evoke emotions [68, 1], and movie genres such as *thriller*, *comedy* or *horror* are explicitly defined by the emotions they evoke. Also, different from existing affective databases, ASCERTAIN comprises data recorded exclusively using commercial sensors to ensure ecological validity and scalability of our framework for profiling applications.

From the ASCERTAIN data, we first examine correlations among users' valence (V) and arousal (A) self-ratings and their personality dimensions. We then attempt to isolate physiological correlates of emotion and personality. Our analyses suggest that the relationships between emotional states and personality traits are better captured by non-linear rather than linear statistics. Finally, we present single-trial (binary) recognition of A,V and the big-five traits considering physiological responses observed over (a) *all*, and (b) *emotionally homogeneous* (e.g, high A, high V) clips. Superior personality recognition is achieved for (b), implying that personality differences are better revealed on comparing responses to emotionally similar stimuli. The salient aspects of ASCERTAIN are:

1. To our knowledge, ASCERTAIN is the first physiological database that facilitates both emotion and personality recognition. In social psychology, personality traits are commonly modeled via questionnaires or social behavioral cues. Instead, this is one of the first works to assess personality traits via affective physiological responses (the only other work similar to ours is [1]).

2. Different from popular recent affective computing databases such as the DEAP [89], MAHNOB [149] and DECAF [3] databases, we use *wearable*, *off-the-shelf* sensors for physiological recordings. This enhances the ecological validity of the ASCERTAIN framework, and above-chance

recognition of emotion and personality confirms its utility and promise for commercial applications.

3. We present interesting insights concerning correlations among affective and personality attributes. Our analyses suggest that the emotion–personality relationship is better captured via non-linear statistics. Also, personality differences are better revealed on comparing user responses to emotionally similar videos (or more generally, under similar affect inducement).

The chapter is organized as follows– Section 2.2 reviews related literature to motivate the need for ASCERTAIN, while Section 2.3 explains the materials and methods employed for data compilation. Section 2.4 presents descriptive statistics, while correlations among users' affective ratings and personality dimensions are analyzed in Section 2.5. Section 2.6 details physiological correlates of emotion and personality, while Section 2.7 presents recognition experiments. Section 2.8 discusses the correlation and recognition results, and Section 2.9 concludes the chapter with key observations.

2.2 Related Work

This section reviews related work focusing on (a) multimodal affect recognition, (b) personality assessment and (c) the personality–affect relationship.

Emotions are an important factor in human life. Any behaviour and environmental stimulus may cause these psychological effects in us and influences our interpretation of the environment and our consequent behaviour. Emotional intelligence, the capability to recognize and understand affective states of the person we are communicating with, is of great importance for our success in social interaction and therefore in life in general. The capability of understanding emotions that comes naturally to most people is very difficult to implement in the current intelligent systems [131]. During the last decades emotions became an increasingly interesting topic in research. So far many studies investigated how emotions are induced through speech and several algorithms were developed to analyse emotional content in speech [48] and were integrated in speech recognition systems e.g. in robots [137]. Also the use of images to induce emotions has been widely studied [143] and research investigating the emotional responses to videos are steadily increasing.

2.2.1 Emotion Theories

In order to understand the phenomenon of feelings, emotion theories have been developed to build hypotheses and run experiments. Mainly two concepts are used today. One classifies emotions in six distinct universal groups [47]: happiness, sadness, anger, fear, surprise and disgust. Each of these basic emotions has a unique facial expression that can be easily detected [53]. For continuous measurements it is more common to use the dimensional concept of emotion. Already in the late 19th century Wundt [185] proposed that emotions could be classified along three dimensions: pleasure, arousal and dominance. To assess the emotions evoked by different stimuli often a technique based on the model of Russell is used [141]. Using this, subjects have to generate 18 ratings for bipolar adjective pairs (e.g. bored-relaxed). To save time and effort the subjects can directly rate arousal (calm-excited) and the valence (unpleasant-pleasant) leading to the same results [21] which can be displayed on a two-dimensional plane with the two axes valence (miserable-pleased) and arousal (sleepiness-aroused).

2.2.2 Physiological responses

To capture stress and emotions some studies have been done with multi-modal emotion recognition systems using facial expressions [145]. Even if facial expressions are closely correlated to emotions they can be controlled consciously and are therefore not completely reliable. Additionally, the face mainly displays the conscious aspect of an emotion. Physiological reactions caused by the peripheral nervous system are unconscious and cannot easily be controlled willingly. Signals that effectively encode emotions are for example: Electrocardiogram (ECG) measuring the heart activity, Electroencephalogram (EEG) measuring the electrical activity of the brain through the scalp, and galvanic skin conductance (GSR) measuring the electrical conductance of the skin [96]. Lisetti and Nasoz [105] review the technologies for capturing physiological signals that are associated with emotions using wearable devices. They use GSR, heart rate and temperature to predict emotions with different algorithms.

2.2.3 Multimodal affect recognition

As emotions are conveyed by content creators using multiple means (audio, video), and expressed by humans in a number of ways (facial expressions, speech and physiological responses), many affect recognition (AR) methods employ a multimodal framework. Common content-based modalities employed for AR include audio [14, 15, 98], visual [109, 191, 132] and audio-visual [30, 60, 144]. Recent AR methodologies have focused on monitoring user behavior via the use of physiological sensors (see [177] for a review). Emotions induced by music clips are recognized via heart rate, muscle movements, skin conductivity and respiration changes in [86]. Lisetti and Nasoz [105] use GSR, heart rate and temperature signals to recognize emotional states. As part of the HUMAINE project [46], three naturalistic and six induced affective databases containing multimodal data (including physiological signals) are compiled from 8–125 participants.

Koelstra *et al.* [91] analyze blood volume pressure, respiration rate, skin temperature and Electrooculogram (EOG) patterns for recognizing emotional states induced by 40 music videos. MAHNOB-HCI [149] is a multimodal database containing synchronized face video, speech, eye-gaze and physiological recordings from 27 users. Abadi *et al.* [1] study Magnetoencephalogram (MEG), Electromyogram (EMG), EOG and ECG responses from users for music and movie clips, and conclude that better emotion elicitation and AR are achieved with movie clips.

2.2.4 Personality recognition

The ***big-five*** or five-factor model [40] describes human personality in terms of five dimensions– Extraversion (*sociable* vs *reserved*), Neuroticism or the degree of emotional stability (*nervous* vs *confident*), Agreeableness (*compassionate* vs *dispassionate*), Conscientiousness (*dutiful* vs *easy-going*) and Openness (*curious/creative* vs *cautious/conservative*).

A comprehensive survey of personality computing approaches is presented in [171]. The traditional means to model personality traits are questionnaires or self-reports. Argamon *et al.* [7] use lexical cues from informal texts for recognizing Extraversion (*Ex)* and Neuroticism (*Neu*). Olguin *et al.* [129] and Pineda *et al.* [5] show that non-verbal behavioral measures acquired using a sociometric badge such as the amount of speech and physical activity, number of face-to-face interactions and physical proximity to other objects is highly correlated with personality. Much work has since employed non-verbal behavioral cues in social settings for personality recognition including [100], where *Ex* is recognized using speech and social attention cues in round-table meetings, while [159, 192] predict *Ex* and *Neu* from proxemic and attention cues in party settings.

Among works that have attempted recognition of all five personality factors, Mairesse *et al.* [112] use acoustic and lexical features, while Staiano *et al.* [153] analyze structural features of individuals' social networks. Srivastava *et al.* [152] automatically complete personality questionnaires for 50 movie characters utilizing lexical, audio and visual behavioral cues. Brouwer *et al.* [22] estimate personality traits via physiological measures, which are revealed subconsciously and more genuinely (less prone to manipulation) than questionnaire answers. In a gaming-based study, they observe a negative correlation between (i) heart rate and *Ex*, and (ii) skin-conductance and *Neu*.

2.2.5 Personality-Affect relationship

The relationship between personality and affect has been extensively examined in social psychology [183], but not in a computational setting. Eysenck's seminal personality theory [52] posits that extraverts require more external stimulation than introverts, and that neurotics are aroused more easily. Many studies have since studied the personality–affect relationship by examining explicit or implicit user responses. Personality effects on brain activation related to valence (V) and arousal (A) is investigated in [85], which concludes that *Neu* correlates negatively with positive V, and positively with A. In an EEG-based study [154], a negative correlation is observed between *Ex* and A, while a positive correlation is noted between *Neu* and A especially for negative valence stimuli.

The impact of personality traits on affective user ratings is studied using path analysis in [163]. Feedback scores from 133 students are analyzed in [127] to conclude that neurotics experience positive emotions similar to emotionally stable counterparts in pleasant situations, even though they may experience negative emotions more strongly. Event-related potentials and heart rate changes are studied in [44] to confirm a positive correlation between *Neu* and A for negative stimuli, while a signal-detection task is used in [70] to suggest that extraverts are generally less aroused than introverts. Brumbaugh *et al.* [27] examine correlations among the big-five traits, and find *Ex* and *Neu* to be associated with increased A while viewing negative videos. Abadi *et al.*[3] attempt recognition of the big-five traits from affective physiological responses, and our work is most similar to theirs in this respect. Nevertheless, we consider user responses to a larger stimulus set in this work (36 clips vs 16 in [3]), and show superior personality trait recognition on comparing physiological responses to emotionally homogeneous clips.

2.2.6 Spotting the research gap

Examination of related literature reveals that AR methodologies are increasingly becoming *user-centric* instead of *content-centric*, suggesting that emotions better manifest via human behavioral cues rather than multimedia content-based (typically audio, visual and speech-based) cues. Nevertheless, the influence of psychological factors such as personality on emotional behavior has hardly been examined, in spite of prior work suggesting that personality affects one's (i) feelings [183, 104], (ii) emotional perception [85, 154] and (iii) multimedia preferences [93, 146].

Motivated by the above findings and the lack of publicly available data sets positioned at the intersection of personality and affect, we introduce ASCERTAIN, a multimodal corpus containing physiological recordings of users viewing emotional videos. ASCERTAIN allows for inferring both personality traits and emotional states from physiological signals. We record GSR, EEG, ECG signals using wearable sensors, and facial landmark trajectories (EMO) using a web-camera. In the light of recent technological developments, these signals can be acquired and analyzed instantaneously. Also, Wang and Ji [177] advocate the need for less intrusive sensors to elicit natural emotional behavior from users. Use of wearable sensors is critical to ensure the ecological validity, repeatability and scalability of affective computing studies, which are typically conducted in controlled lab conditions and with small user groups.

Table 2.2 presents an overview of publicly available user-centric AR datasets. Closest to ASCERTAIN is the dataset of Abadi et al. [3], in which the authors use both movie and music clips to elicit emotions. In contrast, we record our signals using affordable sensors and provide personality annotations in addition to emotional responses. Apart from being one of the largest datasets in terms of the number of participants and stimuli examined for analysis, ASCERTAIN

Table 2.2: Comparison of user-centered affective databases. We point out the number of subjects (N), the number of stimuli, the type of recorded signals and the annotations collected. 'var' denotes variable.

Name	**N**	**Stimuli**	**Recorded signals**	**Annotations**		**Comments**
				Affect	**Personality**	
HUMAINE [46]	var	var	audio, visual, physiological	yes	no	includes 6 sub-collections (some non-public)
DEAP [89]	32	40	physiological	yes	no	focus on music videos
DECAF [3]	30	76	face, physiological	yes	no	compares music and movie clips
MAHNOB-HCI [149]	27	20	face, audio, eye gaze, physiological	yes	no	includes video and image stimuli
ASCERTAIN	36	36	face, physiological	yes	yes	connects emotion and personality

is also the first database to facilitate the investigation of the personality–affect relationship.

2.3 ASCERTAIN Overview

Figure 2.1 presents an overview of the ASCERTAIN framework and a summary of the compiled data is provided in Table 2.1. To study the personality–affect relationship, we recorded users' physiological responses as they viewed the affective movie clips used in [2]. Additionally, their explicit feedback, in the form of *arousal*, *valence*, *liking*, *engagement* and *familiarity* ratings, were obtained on viewing each clip. Finally, personality measures for the big-five dimensions were also compiled using a big-five marker scale (BFMS) questionnaire [133]. We now describe (1) the procedure adopted to compile users' emotional ratings, personality measures and physiological responses, and (2) the physiological features extracted to measure users' emotional responses.

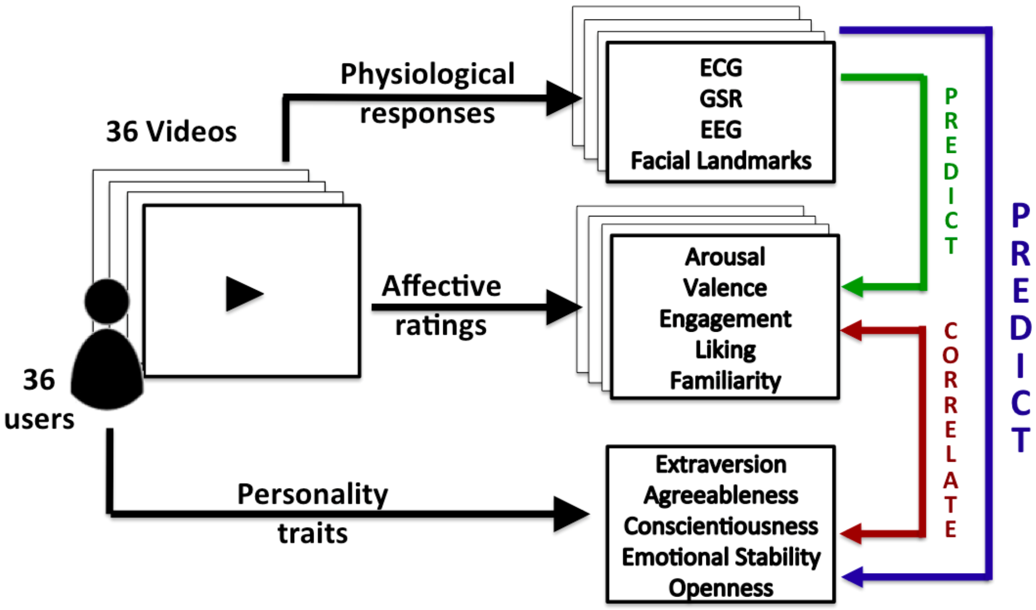


Figure 2.1: ASCERTAIN study overview.

2.3.1 Materials and Methods

Subjects: 36 university students (12 female, mean age = 29.2) from various countries participated in the study. All subjects were fluent in English and were habitual Hollywood movie watchers.

Materials: One PC with two monitors was used for the experiment. One monitor was used for video clip presentation at 1024 × 768 pixel resolution with 60 Hz screen refresh rate, and was placed roughly one meter before the user. The other monitor allowed the experimenter to verify the recorded sensor data. Following informed consent, physiological sensors were positioned on the user's body as shown in Figure 2.2(a). The GSR sensor was tied to the left wrist, and two electrodes were fixed to the index and middle finger phalanges. Two measuring electrodes for ECG were placed at each arm crook, with the reference electrode placed at the left foot. A single dry-electrode EEG device was placed on the head like a normal headset, with the EEG sensor touching the forehead and the reference electrode clipped to the left ear. EEG data samples were logged using the *Lucid Scribe* software, and all sensor data were recorded via bluetooth. A webcam was used to record facial activity. Synchronized data recording and pre-processing were performed using MATLAB Psychtoolbox[3].

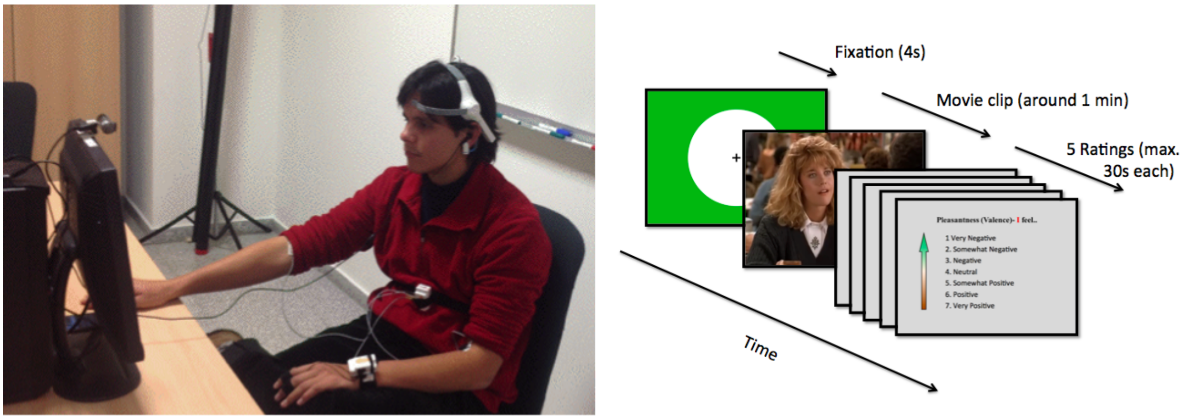


Figure 2.2: Participant with sensors (EEG, ECG and GSR visible) during the experiment (left) and timeline for each trial (right).

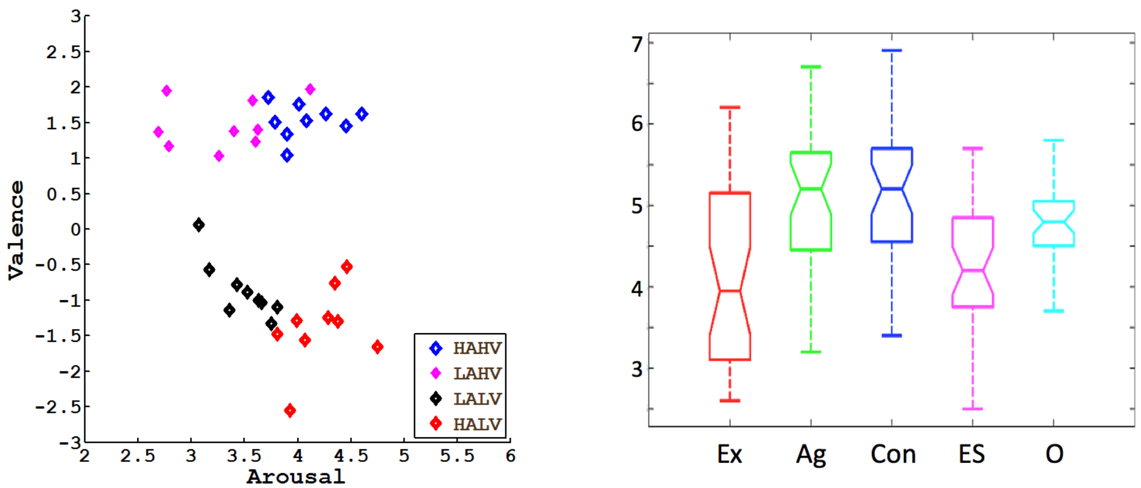
Protocol: Each user performed the experiment in a session lasting about 90 minutes. Viewing of each movie clip is denoted as a *trial*. After two practice trials involving clips that were not part of the actual study, users watched movie clips randomly shown in two blocks of 18 trials, with a short break in-between to avoid fatigue. In each trial (Figure 2.2b), a fixation cross was displayed for four seconds followed by clip presentation. On viewing each clip, users self-reported their emotional state in the form of affective ratings within a time limit of 30 seconds. They also completed a personality questionnaire after the experiment.

Stimuli: We adopted the 36 movie clips used in [2] for our study. These clips are between 51–127 sec long (μ =80, σ =20), and are shown to be uniformly distributed (9 clips per quadrant) over the arousal-valence (AV) plane.

[3]http://psychtoolbox.org/

Affective ratings: For each movie clip, we compiled valence (V) and arousal (A) ratings reflecting the user's affective impression. A 7-point scale was used with a -3 (*very negative*) to 3 (*very positive*) scale for V, and a 0 (*very boring*) to 6 (*very exciting*) scale for A. Likewise, ratings concerning engagement (*Did not pay attention – Totally attentive*), liking (*I hated it – I loved it*) and familiarity (*Never seen it before – Remember it very well*) were also acquired. Mean user V,A ratings for the 36 clips are plotted in Figure 2.3(b), and are color-coded based on the ground-truth ratings from [2]. Ratings form a 'C'-shape in the AV plane, consistent with prior affective studies [89, 2].

Personality scores: Participants also completed the big-five marker scale (BFMS) questionnaire [133] which has been used in many personality recognition works [192, 100, 159]. Scale distributions for the big-five traits are shown in Figure 2.3(b).



(a) (b)

Figure 2.3: (a) Mean Arousal-Valence (AV) ratings for the 36 movie clips used in our experiment and (b) Box-plots showing distribution of the big-five personality trait scores for 36 users.

2.3.2 Physiological feature extraction

We extracted physiological features corresponding to each trial over the final 50 seconds of stimulus presentation, owing to two reasons: (1) The clips used in [3] are not emotionally homogeneous, but are more emotional towards the end. (2) Some employed features (see Table 2.3) are nonlinear functions of the input signal length, and fixed time-intervals needed to be considered as the movie clips were of varying lengths. Descriptions of the physiological signals examined in this work are as follows.

Table 2.3: Extracted features for each modality (feature dimension stated in parenthesis). *Statistics* denote mean, standard deviation (std), skewness, kurtosis of the raw feature over time, and % of times the feature value is above/below mean±std.

Modality	Extracted features
ECG (32)	Ten low frequency ([0-2.4] Hz) power spectral densities (PSDs), four very slow response ([0-0.04] Hz) PSDs, IBI, HR and HRV statistics.
GSR (31)	Mean skin resistance and mean of derivative, mean differential for negative values only (mean decrease rate during decay time), proportion of negative derivative samples, number of local minima in the GSR signal, average rising time of the GSR signal, spectral power in the [0-2.4] Hz band, zero crossing rate of skin conductance slow response ([0-0.2] Hz), zero crossing rate of skin conductance very slow response ([0-0.08] Hz), mean SCSR and SCVSR peak magnitude
Frontal EEG (88)	Average of first derivative, proportion of negative differential samples, mean number of peaks, mean derivative of the inverse channel signal, average number of peaks in the inverse signal, statistics over each of the 8 signal channels provided by the Neurosky software
EMO (72)	Statistics concerning horizontal and vertical movement of 12 motion units (MUs) specified in [80].

Galvanic Skin Response (GSR): GSR measures transpiration rate of the skin. When two electrodes are positioned on the middle and index finger phalanges and a small current is sent through the body, resistance to current flow changes with the skin transpiration rate. Most of the GSR information is contained in low-frequency components, and the signal is recorded at 100 Hz sampling frequency with a commercial bluetooth sensor. Following [86, 89, 149], we extracted 31 GSR features listed in Table 2.3.

Electroencephalography (EEG): EEG measures small changes in the skull's electrical field produced by neural activity, and information is encoded in the EEG signal amplitude as well as in certain frequency components. We used a commercial, single dry-electrode EEG sensor[4], which records eight information channels sampled at 32 Hz. The eight channels are respectively raw EEG, attention and meditation level, alpha, beta, delta, gamma and theta components. The recorded information includes frontal lobe activity, level of facial activation, eye-blink rate and strength, which are relevant emotional responses.

Electrocardiogram (ECG): Heart rate characteristics have been routinely used for user-centered emotion recognition. We performed R-peak detection on the ECG signal to compute users' inter-beat intervals (IBI), heart rate (HR), and the heart rate variability (HRV). We also extracted power spectral density (PSD) in low frequency bands as in [86, 149].

Facial landmark trajectories (EMO): A facial feature tracker [80] was used to compute displacements of 12 interest points or motion units (MU) in each video frame. We calculated 6 statistical measures for each landmark to obtain a total of 72 features (Table 2.3).

[4] www.neurosky.com

2.3.3 Data Quality

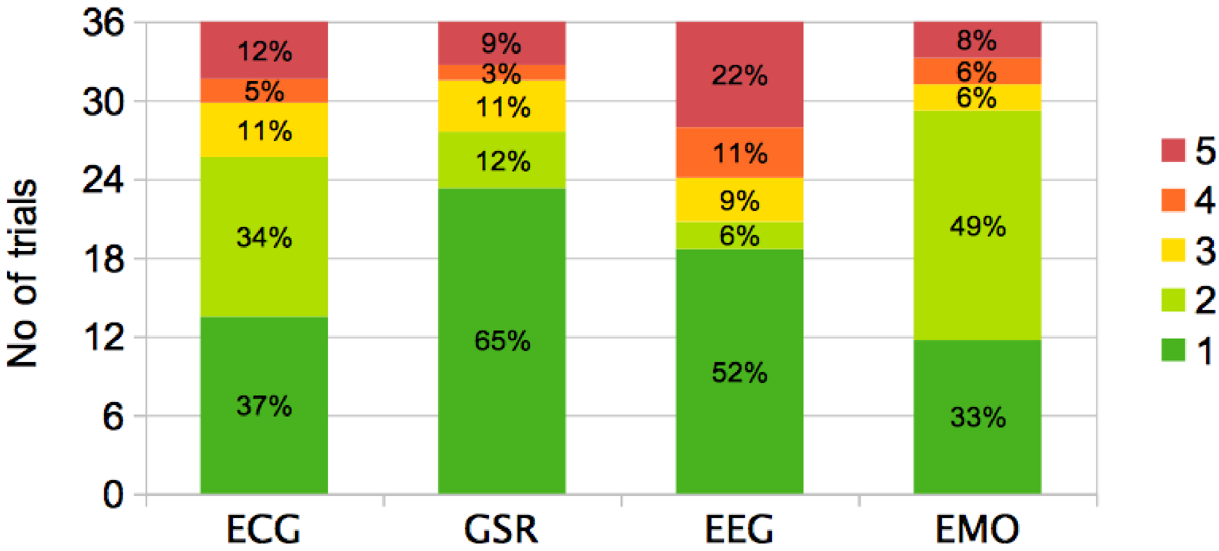


Figure 2.4: Bar plot showing proportion of trials for which data quality ranges from best (1) to worst (5).

A unique aspect of ASCERTAIN with respect to prior affective databases is that physiological signals are recorded using commercial and minimally invasive sensors that allow body movement of participants. However, it is well known that body movements can degrade quality of the recorded data, and such degradation may be difficult to detect using automated methods. Therefore, we plotted the recorded data for each modality and trial, and rated the data quality manually on a scale of 1 (*good data*)–5 (*missing data*). For ECG, we evaluated the raw signal from each arm as well as the R-peak amplitudes. For GSR, we examined the extent of data noise, and rated EEG (i) on the raw signal, (ii) by summarizing the quality of δ ($<$ 4 Hz), θ (4–7 Hz), α (8–15 Hz), β (16–31 Hz) and γ ($>$ 31 Hz) frequency bands, and (iii) on the pre-calculated *attention* and *meditation* channels available as part of the EEG data. Plots and tables with explanations on data quality are available with the dataset. Figure 2.4 presents an overview of the data quality for the four considered modalities, with the proportion of trials for which the quality varies from 1–5 highlighted. About 70%

of the recorded data is good for all modalities except EEG, with the facial video data being the cleanest. Maximum missing data is noted for EEG, reflecting the sensitivity of the EEG device to head movements.

2.4 Descriptive Statistics

In this section, we present statistics relating to user self-reports and personality scores.

2.4.1 Analysis of Self-ratings

As mentioned previously, we selected 36 movie clips such that their emotional ratings were distributed uniformly over the AV plane as per ground-truth ratings in [1], with 9 clips each corresponding to the HAHV (high arousal-high valence), LAHV (low arousal-high valence), LALV (low arousal-low valence) and HALV (high arousal-low valence) quadrants[5]. The targeted affective state was mostly reached during the ASCERTAIN study as shown in Figure 2.3(a). A t-test revealed significantly higher A ratings for HA as compared to LA stimuli ($t(34) = 5.8889, p < 0.0001$). Similarly, V ratings for HV and LV clips were significantly different ($t(34) = 17.9621, p < 0.0001$). Overall, emotion elicitation was more consistent for valence as compared to arousal like also found by prior works [1, 89]. Emotion elicitation was easier for the valence dimension than for the arousal one. This is due to the fact that it is hard to induce strong emotional responses while keeping a low level of arousal. This effect is visible from the C-shape the clip ratings form in the AV plane (Figure 2.3a) being consistent with previous studies [89]. In Figure 2.5 the rating distributions are separated for each quadrant. While HV and LV video ratings are clearly separated, HA and LA have some overlap.

[5]For consistency's sake, quadrant-wise video labels derived based on ratings from [3] are used in this work.

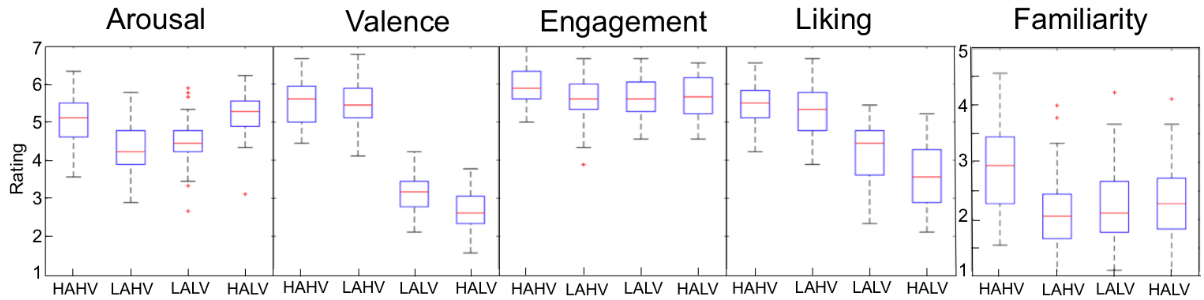


Figure 2.5: Boxplots of the mean Arousal, Valence, Engagement, Liking and Familiarity ratings for the different video sets.

We computed agreement among participants' A,V ratings using the Krippendorff's alpha metric– mean agreement for A and V were respectively found to be 0.11 and 0.61, implying more consensus for clip valence as for arousal. We then computed the agreement between the ASCERTAIN population and the DECAF [3] groundtruth using the Cohen's Kappa measure. To this end, we computed the agreement between ground-truth (GT) labels from [1] and each user's A,V labels assigned as *high*/*low* based on the mean rating– the mean agreement over all users for A and V was found to be 0.22 and 0.72 respectively. Finally, we computed the agreement between GT and the ASCERTAIN population based on the mean A,V rating of all users– here, an agreement of 0.89 was observed for A and perfect agreement (of 1) was noted for V. Overall, these measures suggest that while individual-level differences exist in affective perception of the movie clips, there is high agreement between overall assessments of the ASCERTAIN and DECAF populations implying that the considered movie clips are effective for emotion elicitation.

Figure 2.5 presents box-plots describing the distribution of the arousal (A), valence (V), engagement (E), liking (L) and familiarity (F) user ratings for quadrant-based videos. Clearly, low-arousal videos are perceived as more 'neutral' in terms of A and V, which leads to the 'C'-shaped distribution in Figure 2.3(a). All videos are perceived as sufficiently engaging, while HV clips are more liked than LV clips. Also, the presented movie clips were not very conversant to participants, suggesting that the ASCERTAIN results are overall unlikely to be modulated by familiarity biases.

2.4.2 Correlating Affective Ratings and Personality Scales

To examine relationships between the different user ratings, we computed Pearson correlations among self-reported attributes as shown in Table 2.4. Since the analysis involves attribute ratings provided by 36 users for 36 clips, we accounted for multiple comparisons by limiting the false discovery rate (FDR) to within 5% using the procedure outlined in [16]. Highlighted numbers denote correlations found to be significant over at least 9 users (25% of the population) adopting the above methodology.

Focusing on significant correlations, A is moderately correlated with E and L and there also exists a moderate correlation between E and L. Also, V is found to correlate strongly with L mirroring the observations of Koelstra *et al.* [89]. Correlations between F and E, as well as between F and L confirm the mere exposure effect observed in [20], which attributes liking to familiarity. Nevertheless, different from [89] with music videos where a moderate correlation is noted between A and V ratings, we notice that the A and V dimensions are

Table 2.4: Mean Pearson correlations between self-ratings across users. *s denote significant correlations ($p < 0.05$) upon limiting FDR to 5%.

	A	**V**	**E**	**L**	**F**
Arousal	1	0.04	0.40*	0.21*	0.17
Valence		1	0.21	0.69*	0.18
Engagement			1	0.44*	0.27*
Liking				1	0.37*
Familiarity					1

Table 2.5: Pearson correlations between personality dimensions (* $\Rightarrow p < 0.05$)

	E	**A**	**Co**	**ES**	**O**
Extraversion	1	0.49*	0.13	-0.02	0.45*
Agreeableness		1	0.10	0.34*	0.35*
Conscientiousness			1	0.29	0.15
Emotional Stability				1	-0.07
Openness					1

uncorrelated for the ASCERTAIN study, which again reinforces the utility of movie clips as good control stimuli. To validate our experimental design, we tested for effects of video length on A,V ratings but did not find any.

Table 2.5 presents Pearson correlations between personality dimensions. Again focusing on significant correlations, moderately strong and positive correlations are noted between Extraversion (*Ex*) and Agreeableness (*Ag*), as well as between *Ex* and Openness (*O*)– prior studies have noted that *Ex* and *O* are correlated via the sensation seeking construct [6]. *Ag* is also found to moderately and positively correlate with Emotional Stability (*ES*) and *O*. Conversely, the least correlations are observed between (i) *Ex* and *O*, and (ii) *ES* and *O*.

Partial correlations between self-rated and personality attributes are tabulated in Table 2.6. No significant correlation is noted between personality scales and mean user emotional ratings acquired for all movie clips, but some significant correlates are nevertheless observed when mean ratings for quadrant-wise (or emotionally similar) videos are considered. Focusing on significant correlates, *Ag* is positively correlated with L, but negatively with V ratings for HAHV videos. Surprisingly, a negative correlation is noted between *O* and E for LAHV clips. Consistent with prior studies [39], V is positively correlated with *Ex* in general, with a significant and moderately positive correlation noted for LALV clips. Finally, a moderately negative correlation is observed between mean Valence ratings and *ES* for LALV clips consistent with the observations made in [127].

Table 2.6: Partial correlations between personality scales and self-ratings (* $\Rightarrow p < 0.05$).

		Ex	Ag	Co	ES	O
All	**Arousal**	0.03	-0.19	0.06	-0.15	0.04
	Valence	0.25	-0.06	-0.06	-0.28	0.07
	Engage	-0.25	-0.09	0.15	-0.07	-0.21
	Liking	-0.16	0.23	-0.15	0.30	0.12
HAHV	**Arousal**	-0.01	-0.10	-0.08	-0.03	0.03
	Valence	-0.26	-0.38*	-0.06	-0.07	-0.18
	Engage	-0.13	0.19	0.22	0.03	-0.05
	Liking	0.31	0.35*	0.00	0.12	0.21
LAHV	**Arousal**	0.10	0.01	0.09	-0.19	0.14
	Valence	0.02	0.09	0.01	-0.16	0.06
	Engage	-0.27	0.05	-0.11	-0.09	-0.37*
	Liking	0.12	-0.08	0.07	0.14	0.25
LALV	**Arousal**	0.05	-0.26	0.05	-0.27	0.03
	Valence	0.40*	-0.04	-0.05	-0.39*	0.04
	Engage	-0.10	-0.14	0.09	-0.26	-0.10
	Liking	-0.26	0.18	-0.15	0.30	0.11
HALV	**Arousal**	0.15	-0.15	-0.07	-0.10	-0.03
	Valence	0.25	-0.02	-0.11	-0.19	0.10
	Engage	-0.24	-0.23	0.12	-0.02	-0.17
	Liking	-0.29	0.22	-0.25	0.26	-0.07

We also performed linear regression analyses with user self ratings as predictors and personality attributes as the target variables for the different video sets, and the coefficients of determination/squared correlations (R^2) for the different video sets are presented in Table 2.7. R^2 values with the three best predictors along with the predictor names are listed outside parentheses, while squared correlations with the full model are listed within braces. It is easy to observe from

Table 2.7: R^2 and best three predictors for the five personality dimensions. Full model coefficients are shown in parentheses. * $\Rightarrow p < 0.05$.

	Ex	**Ag**	**Co**	**ES**	**O**
All	0.11 (0.11) V,E,L	0.09 (0.10) A,E,L	0.06 (0.07) V,E,L	0.13 (0.13) A,V,L	0.06 (0.07) V,E,L
HAHV	0.11 (0.11) V,E,L	**0.19*** (0.19) V,E,L	0.05 (0.05) A,V,E	0.01 (0.02) A,V,L	0.05 (0.05) V,E,L
LAHV	0.08 (0.08) A,E,L	0.01 (0.01) V,E,L	0.03 (0.03) A,E,L	0.09 (0.10) A,V,L	**0.20*** (0.20) A,E,L
LALV	**0.20*** (0.20) V,E,L	0.11 (0.11) A,E,L	0.05 (0.05) A,E,L	0.16 (0.21) A,V,L	0.03 (0.03) V,E,L
HALV	0.14 (0.16) V,E,L	0.16 (0.16) A,E,L	0.15 (0.16) V,E,L	0.07 (0.07) A,V,L	0.06 (0.06) V,E,L

the table that (i) there is little difference in the predictive power of the best-three-predictor and full models, and (ii) the linear models have rather limited predictive power, with the best models explaining only 20% of the personality scale variance. Overall, Tables 2.6 and 2.7 cumulatively suggest that the relationship between emotional and personality variables is not well modeled using linear statistics, and it is perhaps worthwhile to explore the use of non-linear measures to this end. Also, given the significance of the relationship between emotional (A,V) attributes and personality dimensions, and the high degree of correlation between E,L and A,V ratings, we will only focus on arousal and valence in the rest of the chapter.

2.4.3 Mutual Information Analysis

Mutual information (MI) is a popular metric to capture non-linear relationships between two random variables, and measures how much information is known

about one variable given the other. Formally, the MI between two random vectors $X = \{x\}$ and $Y = \{y\}$ is defined as:
$MI(X,Y) = \sum_{x,y} P_{XY}(x,y) log \frac{P_{XY}(x,y)}{P_X(x).P_Y(y)}$ where $p_{XY}(x,y)$ is the joint probability distribution, while $P_X(x)$ and $P_Y(y)$ are the respective marginal probabilities. We attempted to describe the relationship between emotional ratings and personality scales via the normalized mutual information (NMI) index [158] defined as: $NMI(X,Y) = \frac{MI(X,Y)}{\sqrt{(H(X)H(Y))}}$, where $H(X)$ and $H(Y)$ denote entropies of X and Y.

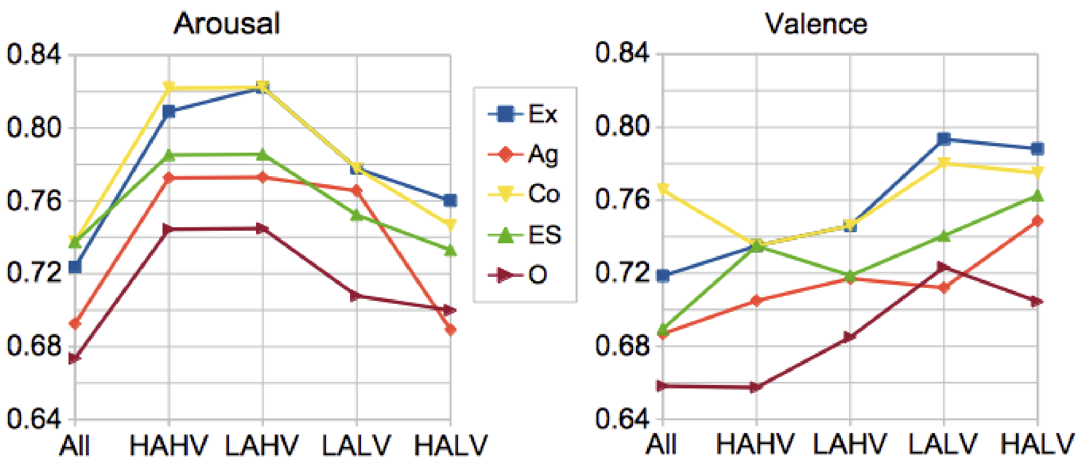


Figure 2.6: NMI between big-five trait scales and A (left), V (right) ratings.

Computed NMI values with personality scales for arousal and valence ratings are shown in Figure 2.6. In contrast to linear measures, both A and V share a high degree of mutual information with all the five personality traits. While considering all the movie clips, MI is generally higher for A as compared to V. Among personality traits, Conscientiousness (*Con*) and *Ex* share the most MI with affective attributes (note that in contrast, little correlation is observed between *Con* and A,V in Table 2.6), while lowest MI is noted for Openness. For both A and V, higher MI with personality dimensions is noted when ratings for quadrant-based videos are considered instead of all movie clips. One notable difference exists between A and V though– higher MI with arousal is noted for high V as compared to low V videos. In contrast, for all personality traits barring *Ag*, greater MI with valence is observed for LV clips than for HV clips.

2.5 Personality measures vs user ratings

We now examine the relationship between user V,A ratings and personality scales in the context of hypotheses (H1–H4) put forth in the literature. To this end, we determined *high*/*low* trait groups (*e.g.*, emotional stable vs neurotic) for each personality dimension by dichotomizing personality measures based on the median score– this generated balanced *high* and *low* sets for three traits, with an imbalanced split (19 vs 17) obtained for Conscientiousness and Openness. We then proceeded to analyze the affective ratings for each group.

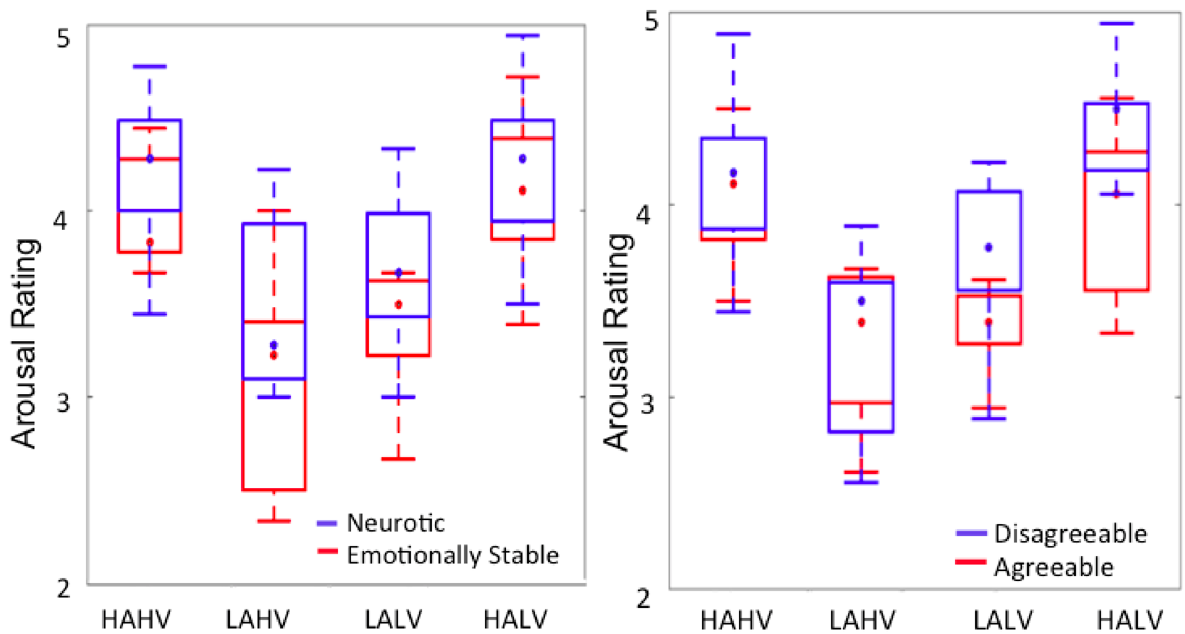


Figure 2.7: Quadrant-wise comparisons of A ratings by Neurotic and Emotionally Stable groups (left), Agreeable and Disagreeable groups (right).

2.5.1 H1: Extraversion vs Arousal and Valence

The correlation between Extraversion and arousal has been investigated in many studies – EEG measurements in [154], signal detection analysis in [70], and fMRI [85] have shown lower arousal in extraverts as compared to introverts, consistent with Eyesenck's personality theory. Also, Extraversion has been found to correlate with positive valence in a number of works [39].

Analyses presented in Table 2.6 reveal a very weak positive correlation between *Ex* and A. While two-tailed t-tests showed that both extraverts and introverts rated high A and low A videos differently ($p < 0.00001$ in both cases), no differences in A ratings could be identified between the two groups considering either all or quadrant-based movie clips. Focusing on V ratings, a generally positive correlation between *Ex* and V can be noted for all video sets with the exception of HAHV clips. A significant positive correlation is however noted only for negative (LALV) videos. Therefore, while statistical analyses do not support the negative correlation between Ex and A, they weakly corroborate the positive correlation between *Ex* and V.

2.5.2 H2: Neuroticism vs Arousal

The relationship between *Neu* and A has been extensively studied and commented on– a positive correlation between *Neu* and A is revealed through fMRI responses [85], and EEG analysis [154] reinforces this observation especially for negative valence stimuli. [127] further remarks that neurotics experience negative emotions stronger than emotionally stable persons.

A ratings provided by the neurotic and *ES* groups were found to be significantly different for low-arousal clips as confirmed by a left-tailed t-test ($t(34) = -1.9555, p = 0.0294$). Quadrant-wise distributions of A ratings for the *ES* and neurotic groups are presented in Figure 2.7 (left). Neurotics are generally found to experience slightly higher arousal than *ES* counterparts. Left-tailed t-tests confirmed that neurotics provided significantly higher A ratings for LALV ($t(16) = -1.7606, p = 0.0487$) clips, and marginally higher A ratings for LAHV ($t(16) = -1.7349, p = 0.0510$) and HAHV ($t(16) = -1.5410, p = 0.0714$) stimuli. In general, our analyses support the observation that Neuroticism is associated with higher A, with the effect being more pronounced for LA videos.

2.5.3 H3: Neuroticism vs Valence

Differing observations have been made regarding the relationship between *Neu* and V. A negative correlation between *Neu* and positive valence is observed in [85], while a positive relationship between the two for low arousal stimuli is noted in [163]. [127] remarks that the *Neu*-V relation is moderated by situation– while neurotics may feel less positive in unpleasant situations, they experience positive emotions as strongly as *ES* subjects in pleasant conditions.

Comparing V ratings of the neurotic and *ES* groups, very similar V ratings are noted for high/low V clips. Quadrant-wise comparisons also failed to reveal any differences. Overall, no definitive relationship was noted between between *Neu* and V.

2.5.4 H4: Openness vs Valence and Arousal

Among the few works to study Openness, [163] notes a positive correlation between Openness and valence under low arousal conditions, which is attributed to the intelligence and sensitivity of creative individuals[6], enabling them to better appreciate subtly emotional stimuli. Right-tailed t-tests to compare V ratings of the *open* and *closed* groups failed to reveal any differences. Quadrant-based comparisons showed that open individuals experienced slightly higher V for LAHV clips ($t(16) = 1.3737, p = 0.0942$).

[6]Creativity strongly correlates with Openness [119].

No significant difference was noted in the A ratings of *open* and *closed* subjects for HA and LA videos. For fine-grained analysis, we again used left-tailed t-tests considering quadrant-wise ratings, which revealed that *closed* individuals experienced slightly higher arousal for HALV clips ($t(16) = -1.3499, p = 0.0979$). In summary, we observed a slightly positive relationship between Openness and A,V as noted in [163].

2.5.5 Agreeableness and Conscientiousness

Quadrant-wise comparison of A ratings by *agreeable* and *disagreeable* groups (Figure 2.7(b)) revealed that *disagreeable* subjects felt more aroused by HALV ($t(16) = -2.0811, p = 0.0269$) and by LALV ($t(16) = -1.8003, p = 0.0453$) clips. The fact that disagreeable persons felt more aroused by low-valence clips could possibly be attributed to their association with negative feelings such as deceit and suspicion.

Conscientiousness scale differences did not significantly influence the VA ratings in any manner.

2.6 Physiological correlates of emotion and personality

From the previous section, we note that the impact of personality differences on the emotion perceived by users is revealed mainly through quadrant-wise comparisons of V and A ratings involving emotionally similar or homogeneous clips. If explicit ratings provided by users are a conscious reflection of their emotional perception, then the analyses employing physiological signals should also reveal similar patterns. We attempt to identify linear and non-linear physiological correlates of emotion and personality considering responses to *all* and *quadrant-specific* clips in this section.

2.6.1 Linear correlates of Emotion and Personality

Table 2.8: Physiological correlates of emotion and personality attributes. R° denotes the number of significant feature correlates, while R^2 is the coefficient of determination for the regression model with the significant correlates as predictors. Bold values denote linear regression models with a significant R^2 statistic.

		Arousal		**Valence**		**Extra.**		**Agreeable**		**Conscient**		**Em. Stab.**		**Open**	
Video Set	Feature	R^o	R^2	R^o	R^2	R^o	R^2	R^o	R^2	R^o	R^2	R^o	R^2	R^o	R^2
All	**ECG**			1	**0.17**	2	0.51					1	0.49	1	0.35
	GSR														
	EMO													1	**0.28**
	EEG	1	0.45	1	0.27	2	0.66	2	0.64	2	**0.68**	2	0.54	2	0.49
HAHV	**ECG**					1	0.46					1	0.45		
	GSR														
	EMO			1	0.59							3	0.75	3	0.77
	EEG			1	0.49	2	0.54							1	0.50
LAHV	**ECG**			1	**0.15**	3	0.55			1	0.44	1	0.46	1	0.42
	GSR														
	EMO	1	**0.20**			1	0.30	1	0.29					2	**0.51**
	EEG	1	0.54											1	0.56
LALV	**ECG**	1	0.41			2	0.55					1	0.44	1	0.42
	GSR														
	EMO			1	0.22	2	0.61	1	0.40					2	0.45
	EEG			1	0.40	1	0.57			3	0.63	1	0.58		
HALV	**ECG**											2	0.57		
	GSR														
	EMO													1	**0.13**
	EEG			1	**0.59**	3	**0.69**			2	0.53			1	0.39

We attempted to discover physiological correlates of emotional attributes and the big-five personality traits via partial Pearson correlations. Given the large number of extracted physiological features (Table 2.3) as compared to the population size for this study, we first performed a principal component analysis (PCA) on each feature modality to avoid overfitting, and retained those components that explained 99% of the variance. This gave us 8–9 predictors for each of the considered modalities. Table 2.8 presents correlations between these principal components, users' affective ratings and personality scales (R°

denotes number of significant correlates). For affective dimensions, we determined significant correlates considering mean user V,A ratings provided for the 36 clips. We also trained regression models with the significantly correlating components as predictors of the dependent emotion/personality variable, and the squared correlations (R^2) of these models are also tabulated.

Examining Table 2.8, the relatively few (maximum of 3) number of significant predictors can be attributed to the sparse number of principal components employed for analysis. Considering correlations with A and V, more correlates are observed for V than for A overall. At least one significant correlate is noted for all modalities except GSR. EEG is the modality found to correlate most with A, with one correlate observed for all and LAHV movie clips. EEG also has the most number of correlates with V (one significant correlate per video set), but this is unsurprising as a number of works have successfully recognized V with commercial-grade sensors [24]. For V, one ECG correlate is noted for all and LAHV videos, and one EMO correlate for HAHV and LALV videos. Prior studies [1, 89] have also noted that these modalities correlate better with V than A.

Focusing on personality dimensions, a larger number of physiological correlates are observed as compared to emotional attributes. Least number of correlates are noted for Agreeableness, while most correlates are noted for Extraversion. The EEG modality again corresponds to the maximum number of correlates, while no correlates are observed for GSR. In general, a larger number of physiological correlates are noted for quadrant-based videos for all traits except Agreeableness. Also, models with a higher R^2 statistic are noted for emotion-wise similar clips, suggesting that the physiology–based linear models are more effective at predicting personality traits while comparing user responses under similar affective conditions. Considering models with significant R^2 values, the highest value of 0.69 is noted for Extraversion for HALV, while none of the linear models can accurately predict Agreeableness and Emotional Stability.

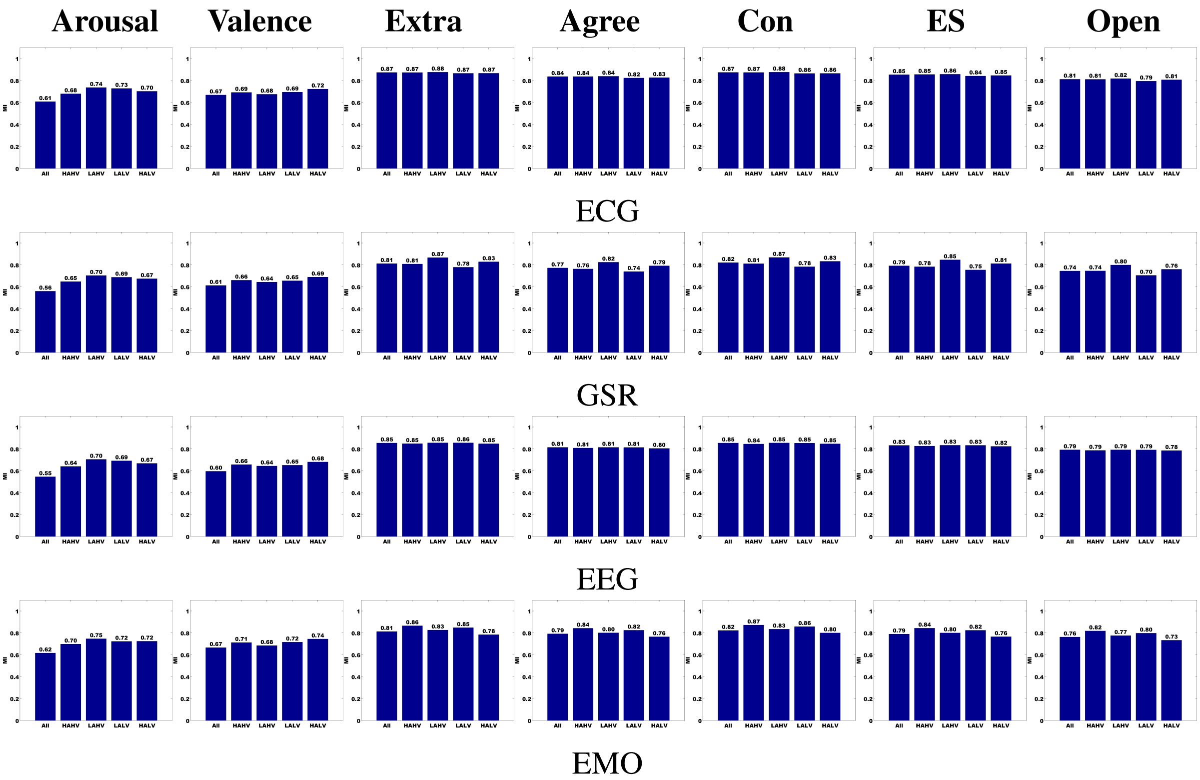


Figure 2.8: (From top to bottom) Bar plots showing the means of the NMI histograms for the four modalities. Best viewed under zoom.

2.6.2 Non-linear correlates

To examine non-linear physiological correlates of emotion and personality attributes, we performed a mutual information analysis between the extracted features from the four modalities and those attributes. Given the varying number of features for each modality, we segregated the NMI distribution over all features and the emotion/personality rating using 10-bin histograms. Figure 2.8 presents the first moment or the mean of the NMI histogram distribution computed over the different video sets for each emotional/personality attribute.

From Figure 2.8 can be noted that personality attributes share more MI with the user's physiological responses than A and V, in line with the linear analyses. Also, for both emotion and personality attributes, the NMI computed for the quadrant-based video sets is generally equal to or higher than the NMI for all movie clips, implying that a fine-grained examination of the relationship between sub-conscious physiological responses and conscious user self-ratings is more informative. Focusing on affective attributes, higher MI between ratings and physiological responses is noted for V considering all movie clips for all modalities, while the highest quadrant-based NMIs are noted with A for all four modalities. Among the four modalities, facial features share the highest MI with both A and V, followed by GSR, ECG and EEG.

Focusing on the big-five personality traits, the highest NMI histogram means over all modalities are observed for Conscientiousness and Extraversion, followed by Emotional Stability, Agreeableness and Openness. This trend is strikingly similar to that obtained from the MI analysis between V,A ratings and personality scales in Section 2.4.3. Examining the sensing modalities, ECG features share the highest MI with personality scales followed by EEG, while EMO and GSR correspond to lower NMI means.

2.7 Recognition results

We performed binary recognition of emotional and personality attributes to evaluate the efficacy of the proposed framework. This section details the experiments and results thereof.

2.7.1 Emotion recognition

A salient aspect of this work is the exclusive use of commercial sensors for examining users' physiological behavior. To evaluate if we can still achieve emotion recognition comparable to prior affective works which use laboratory-grade sensors, we followed a procedure identical to the DEAP study [89]. In particular, the most discriminative physiological features were first identified for each modality using Fisher's linear discriminant with a threshold of 0.3. Features corresponding to each user were then fed to the naive Bayes (NB) and linear SVM classifiers as shown in Table 2.9, with a leave-one-out cross-

Table 2.9: **Affective state recognition** with linear SVM and Naive Bayes (NB) classifiers. Mean F1-scores over all participants for the four modalities, peripheral Signals (ECG + GSR) and late fusion ($\mathrm{W}^{\mathrm{t}}_{\mathrm{est}}$) are shown. Baseline F1-score is 0.5. Maximum unimodal F1-scores are shown in bold.

	ECG		**GSR**		**EMO**		**EEG**		**Peripheral**		$\mathbf{W}^{\mathbf{t}}_{\mathbf{est}}$		**Class Ratio**
	SVM	NB	SVM	NB	SVM	NB	SVM	NB	SVM	NB	SVM	NB	
Arousal	0.54	0.58	0.52	0.58	0.55	0.60	0.59	**0.61**	0.56	0.62	0.62	0.62	0.50
Valence	0.55	0.58	0.52	0.57	0.58	**0.62**	0.60	**0.62**	0.55	0.62	0.64	0.64	0.50

validation scheme employed where one video is held out for testing, while the other videos are used for training. The best misclassification cost parameter C for linear SVM is determined via grid search over $[10^{-3}, 10^{3}]$ again using leave-one-out cross-validation.

Table 2.9 presents the mean F1-scores over all users obtained using the NB and SVM classifiers with unimodal features and the decision fusion ($\mathrm{W}^{\mathrm{t}}_{\mathrm{est}}$) technique described in [89]. In decision fusion,the test sample label is computed as $\sum_{i=1}^{4} \alpha_i^* t_i p_i$. Here, i indexes the four modalities used in this work, p_i's denote posterior SVM probabilities, $\{\alpha_i^*\}$ are the optimal weights maximizing the F1-score on the training set and $t_i = \alpha_i F_i / \sum_{i=1}^{4} \alpha_i F_i$, where F_i denotes the F1-score obtained on the training set with the i^{th} modality. Note from Section 2.3 that there is an equal distribution of high/low A and V, implying a class ratio (and consequently, a baseline F1-score) of 0.5

Observing Table 2.9, above-chance emotion recognition is achievable using the physiological features extracted using commercial sensors. The obtained F1-scores are very similar to DEAP [89], which employs music video excerpts to induce emotions. EEG features produce the best recognition performance for A, while both EEG and facial features produce the best recognition for V. GSR produces the worst recognition performance, and the NB classifier outperforms linear SVM for all the considered features. Finally, the best fusion-based recognition performance of 0.64 is noted for V, and better (unimodal as well as multimodal) recognition is generally noted for valence as in [89, 1].

2.7.2 Personality recognition

For binary personality trait recognition, we first dichotomized the big-five personality trait scores based on the median. This resulted in an even distribution of *high* and *low* trait labels for three traits, while an inexact split (19 vs 17) was obtained for Conscientiousness and Openness. As baselines, we consider majority-based voting and random voting according to class ratio. Based on majority voting, F1-score for the *Con* and *O* traits is 0.35 and 0.33 for the others. For class-ratio based voting, a baseline score of 0.5 is achieved for all traits. We performed PCA on each feature modality in a similar way to linear correlation analyses prior to classification. A leave one-subject-out cross-validation scheme was used to compute the recognition results. Three classifiers were employed for recognition, i) naive Bayes, ii) linear (Lin) SVM and iii) Radial Basis Function (RBF) SVM. The C (linear and RBF SVM) and γ (RBF SVM) parameters were tuned via leave-one-subject-out grid search cross-validation on the training set.

Table 2.10 presents the recognition results, with the best F1-scores achieved using unimodal features denoted in bold. For each personality trait, feature and video set, a better-than-chance recognition F1-score (> 0.5) is achieved with at least one of the employed classifiers. Considering user physiological responses to all affective videos, *Con* is the best recognized personality trait, while *Ag* corresponds to the lowest F1-score. Also, higher recognition scores are generally achieved considering user responses to quadrant-wise videos, in line with the observations from linear and non-linear correlation analyses.

Considering feature modalities, EEG is found to produce the best recognition performance across personality traits and video sets, followed by ECG, GSR and EMO. EEG is found to be the optimal feature for recognizing *ES*, while GSR is the best feature for *Con* recognition. Focusing on classifiers, the non-linear RBF SVM produces the best F1-score for all traits except *Ag* while

Table 2.10: Personality recognition considering affective responses to a) all, and b) emotionally homogeneous stimuli. Maximum F1-scores with unimodal classifiers are shown in bold.

		Extravert			**Agreeable**			**Conscient**			**Em. Stab**			**Open**		
Videos	**Method**	**NB**	**SVM (lin)**	**SVM (rbf)**	**NB**	**SVM (lin)**	**SVM (rbf)**	**NB**	**SVM (lin)**	**SVM (rbf)**	**NB**	**SVM (lin)**	**SVM (rbf)**	**NB**	**SVM (lin)**	**SVM (rbf)**
All	**ECG**	0.45	0.61	0.61	0.49	0.00	0.38	0.44	0.14	0.60	0.52	0.00	0.31	0.57	0.46	0.56
	EEG	0.60	0.59	**0.63**	0.51	0.19	0.15	0.52	0.35	0.51	0.58	0.61	**0.61**	0.55	0.55	0.39
	EMO	0.32	0.41	0.40	**0.60**	0.41	0.42	0.50	0.25	0.15	0.46	0.35	0.35	0.28	0.48	**0.62**
	GSR	0.08	0.00	0.47	0.20	0.00	0.53	0.59	0.60	**0.66**	0.58	0.54	0.55	0.35	0.29	0.61
	$\mathbf{W^t_{est}}$	0.44	0.67	0.61	0.58	0.39	0.49	0.53	0.38	0.60	0.68	0.66	0.66	0.58	0.66	0.66
HAHV	**ECG**	0.24	0.11	**0.67**	0.31	0.00	**0.69**	0.39	0.22	0.22	0.24	0.08	0.19	0.45	0.44	**0.63**
	EEG	0.40	0.29	0.31	0.60	0.23	0.51	0.40	0.28	0.49	0.54	**0.56**	**0.56**	0.51	0.35	0.31
	EMO	0.56	0.42	0.65	0.56	0.33	0.50	0.52	0.38	0.38	0.34	0.17	0.08	0.53	0.41	0.57
	GSR	0.44	0.00	0.50	0.00	0.00	0.36	**0.67**	0.33	0.46	0.54	0.53	0.47	0.36	0.35	0.56
	$\mathbf{W^t_{est}}$	0.57	0.33	0.69	0.58	0.32	0.58	0.67	0.29	0.50	0.58	0.62	0.58	0.58	0.44	0.61
LAHV	**ECG**	0.31	0.03	0.16	0.44	0.03	0.00	0.52	0.61	0.64	0.32	0.00	0.16	0.58	**0.75**	0.69
	EEG	**0.63**	0.60	0.43	**0.51**	0.20	0.39	0.10	0.35	0.27	0.60	**0.62**	0.46	0.60	0.65	0.56
	EMO	0.55	0.35	0.49	0.48	0.31	0.29	0.53	0.60	0.54	0.47	0.23	0.21	0.62	0.00	0.53
	GSR	0.35	0.00	0.08	0.40	0.00	0.44	0.64	0.66	**0.69**	0.61	0.57	0.54	0.29	0.29	0.54
	$\mathbf{W^t_{est}}$	0.42	0.44	0.53	0.61	0.29	0.41	0.69	0.67	0.69	0.55	0.64	0.58	0.67	0.75	0.69
LALV	**ECG**	0.33	0.50	0.31	0.58	0.03	0.53	0.55	0.16	0.24	0.44	0.20	0.50	0.44	0.32	0.33
	EEG	0.60	0.58	0.51	0.42	0.37	0.39	0.29	0.33	0.31	**0.71**	0.52	0.51	**0.67**	0.62	0.65
	EMO	**0.65**	0.23	0.35	0.45	**0.69**	0.66	0.38	0.29	0.13	0.36	0.55	0.38	0.44	0.11	0.03
	GSR	0.19	0.00	0.62	0.30	0.00	0.05	0.61	**0.66**	0.50	0.58	0.61	0.62	0.29	0.29	0.44
	$\mathbf{W^t_{est}}$	0.60	0.58	0.67	0.58	0.71	0.69	0.67	0.69	0.50	0.69	0.61	0.72	0.58	0.64	0.69
HALV	**ECG**	0.41	**0.64**	0.61	0.34	0.00	**0.63**	0.44	0.22	0.18	0.52	0.03	0.72	0.44	0.68	0.69
	EEG	0.59	0.54	0.48	0.45	0.31	0.31	0.57	0.35	0.60	**0.67**	**0.67**	0.49	0.68	0.66	**0.71**
	EMO	0.47	0.35	0.49	0.27	0.35	0.24	0.26	0.35	0.55	0.41	0.35	0.27	0.48	0.33	0.31
	GSR	0.18	0.00	0.36	0.18	0.00	0.18	0.55	**0.64**	0.53	0.58	0.52	0.57	0.36	0.29	0.16
	$\mathbf{W^t_{est}}$	0.60	0.66	0.62	0.41	0.41	0.55	0.61	0.68	0.61	0.68	0.65	0.53	0.68	0.72	0.75

considering user responses to all videos. RBF SVM also produces the best recognition performance in 11 out of 25 (5 personality traits $\times$ 5 video sets) conditions. Linear classifiers perform best for the *ES* trait, producing the best F1-scores for quadrant-specific video sets. Among the two linear classifiers, NB slightly outperforms linear SVM, producing highest F1-scores in eight conditions (as against 7 with Lin SVM).

Fusion-based recognition is beneficial in general, and higher recognition scores as compared to unimodal recognition are achieved. With user responses acquired for all videos, the highest and least fusion-based F1 scores are achieved for the *ES* (0.68 with NB classifier) and *Ag* (0.58 with NB) traits respectively. With quadrant-based videos, a maximum F1-score of 0.75 is noted for Openness (with RBF classifier). For fusion-based recognition, RBF SVM again outperforms the NB and linear SVM classifiers for the vast majority of conditions, and most notably for the Openness trait where it achieves optimal recognition for four of the five video sets.

2.8 Discussion

The correlation analyses and recognition results convey two aspects related to personality recognition from physiological data: (i) A fine-grained analysis of users' physiological responses to quadrant-wise or emotionally similar movie clips enables better characterization of personality differences. This is notable from the increased number of linear physiological correlates for some of the AV quadrants in Table 2.8 for all traits excepting *Ag*, and the generally higher NMI means for quadrant-specific videos in Figure 3.3. Also, higher F1-scores are obtained when physiological responses to emotionally similar clips are used for personality trait recognition. (ii) The relationship between personality scales and physiological features is better captured via non-linear metrics. Considerably high NMI means are noted for all feature modalities and personality traits, and the general effectiveness of the RBF SVM classifier support this observation.

Feature-based relationships are also evident between the correlation and recognition experiments. Given that linear classifiers are employed for affect recognition, it is interesting to note the similarities between Tables 2.8 and 2.9. The most number of affective correlates are noted for EEG and EMO in Table 2.8, and the best recognition performance is achieved with these features (Table 2.9). Conversely, no GSR correlate is observed for A and V, and GSR also achieves the worst emotion recognition performance. For personality recognition, ECG and EEG are found to be the better features from linear and non-linear correlation analyses, and these two modalities also produce better recognition performance. On the other hand, GSR features produce good recognition only for the *Con* trait, with which they share a high degree of MI.

Focusing on personality traits, the least number of linear physiological correlates in Table 2.10 are observed for the *Ag* trait and correspondingly, the worst F1-scores with linear classifiers (especially with linear SVM) are noted for this trait. Linear classifiers perform somewhat better for the *Con* and *Ex* traits, and produce the best recognition performance for the *O* and *ES* traits, for both of which a good number of linear correlates are obtained.

It is appropriate to point out some limitations of this study in general. Weak linear correlations are noted between emotional and personality scores in Table 2.6, and only few physiological correlates of emotion and personality are observed in Table 2.8, which can partly be attributed to the low variance for some of the personality dimensions as seen in Figure 2.3(b). While median-based dichotomization of the personality scores for binary recognition is common in personality research [171], it may not be the most appropriate method. Most user-centered affective studies have also demonstrated recognition in a similar manner and on data compiled from small user populations, due to the general difficulty in conducting large-scale affective experiments. Overall, the general consistency in results from the correlation and recognition experiments suggest that data artifacts have only minimally influenced our analyses, and that reliable affect and personality recognition is achievable via the extracted physiological features. Furthermore, we make the collected data publicly available in order to facilitate related research.

Even though not analyzed in this work, the ASCERTAIN database also includes Familiarity and Liking ratings, which could be useful for other research studies. With familiarity ratings we collected data about how well-known the shown scenes are to the users. Potentially, people with specific personality traits have different physiological responses not only based on the emotional value of a video but also depending on its familiarity to them. Whether people prefer specific videos, depending on their personality type, is relevant for building recommender systems that are personality aware. Such systems are built more frequently nowadays and are appreciated by users as easier to use than common ones based on ratings [78]. Personalized video recommender systems could learn and adapt to users over time by taking physiological signals into account.

Familiarity and Liking ratings could be also used to replicate and extend related studies reported in the literature. For example, the study presented in [184] notes a connection between familiarity, liking and the amount of smiling while listening to music. Also, Hamlen and Shuell [73] find a positive correlation between liking and familiarity for classical music excerpts, which increases when an associated video was presented with the audio. Similar effects could be tested with emotional videos with ASCERTAIN.

Finally, the importance of using less-intrusive sensors for affective computing has also been widely acknowledged [105, 177]. Collecting data using minimally invasive and wearable sensors enables naturalistic user response to the presented stimuli, alleviating the stress caused by cumbersome clinical equipment. Choosing minimally invasive sensors is especially critical when complex behavioral phenomena such as emotions are the subject of investigation. While most of the currently available affective datasets have been compiled using lab equipment [177], ASCERTAIN represents one of the first initiatives to exclusively use wearable sensors for data collection, which enhances its ecological validity.

2.9 Conclusion

This chapter presents ASCERTAIN – a new multimodal affective database compiled using only commercial and wearable sensors. ASCERTAIN contains implicit physiological responses of 36 users collected via EEG, ECG, GSR sensors and a webcam while viewing emotional movie clips, along with their explicit affective ratings and big-five personality trait scores. Among affective datasets, ASCERTAIN is the first that facilitates analysing the relationship between physiological signals, emotional states and personality traits.

Using the collected data, we analyzed the correlations between affective ratings, personality scales and extracted physiological features. We only obtained weak correlations and regression models for emotional ratings and personality scales, but we noted a high degree of mutual information between the two. This implies that the personality–affect relationship is better characterized via non-linear statistics. When including physiological features in the analyses, we found only few linear correlates with affective ratings and personality traits. In contrast we obtained considerably high NMI means over all feature modalities and personality dimensions– also, personality attributes share more MI with physiological signals than arousal and valence share with them. Finally, we achieved emotion recognition performance comparable with former works employing lab-grade sensors. We obtained above-chance personality trait recognition with each of the considered modalities. Personality differences are better revealed when analyzing responses to emotionally similar movie clips, as suggested by both correlation and recognition experiments. Finally, RBF SVM produces the best recognition performance overall among the three classifiers, further supporting a non-linear relationship between emotional responses and personality scales, while EEG and ECG are found to be the two best modalities for personality trait recognition.

We believe that ASCERTAIN will facilitate future affective studies, and encourage further investigation on the relationship between personality and affect, particularly in computational settings. The fact that personality differences are observable from user responses to emotion-wise similar stimuli can pave the way for simultaneous emotion and personality profiling. Our future research will focus on the development of real-time emotion and personality recognition systems, coupled with a deeper examination on the relationship between personality and affective behavior. We will also study how prior knowledge of personality can impact the design of user-centered affective studies.

Chapter 3

Inference of Personality Traits and Affect Schedule by Analysis of Spontaneous Reactions to Affective Videos

This chapter[1] presents a method for inferring the Positive and Negative Affect Schedule (PANAS) and the big-five personality traits of 35 participants through the analysis of their implicit responses to 16 emotional videos. The employed modalities to record the implicit responses are (i) EEG, (ii) peripheral physiological signals (ECG, GSR), and (iii) facial landmark trajectories. The predictions of personality traits/PANAS are done using linear regression models that are trained independently on each modality. The main findings of this study are that: (i) PANAS and personality traits of individuals can be predicted based on the users' implicit responses to affective video content, (ii) ECG+GSR signals yield 70%±8% F1-score on the distinction between extroverts/introverts, (iii) EEG signals yield 69%±6% F1-score on the distinction between creative/non creative people, and finally (iv) for the prediction of agreeableness, emotional stability, and baseline affective states we achieved significantly higher than chance-level results.

[1]This chapter is based on: Mojtaba Khomami Abadi, Juan Abdón Miranda Correa, Julia Wache, Heng Yang, Ioannis Patras, and Nicu Sebe. Inference of personality traits and affect schedule by analysis of spontaneous reactions to affective videos. In 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, 2015. [1]

3.1 Introduction

In human computer interactions, the emotional state of a user is a great source of information to enrich the experience. For instance, in an e-learning scenario the computer may adapt the content shown to the user depending on how easy (boring) or difficult (stressing) the content is perceived by the user. Recognizing the emotional state of the user to enhance the user experience has been targeted intensively in affective computing research [145]. Some studies used the explicit responses of the users (e.g. interrupting him/her and asking to self-assess his/her emotional state) to access their emotions. However, most of the recent studies [105] try to avoid interruptions and instead analyze the implicit emotional responses of users (e.g. facial expressions, physiological signals) to automatically infer their emotional state. Emotional responses of humans are influenced by some factors (such as mood, baseline affective schedule, personality, temper and memories), that make the emotion recognition tasks more user-specific. However, by learning the effect of the factors on the emotional behaviors, we can generalize the user-specific model to cross-users models.

The objective of this work is to study the relation between these factors, in particular personality and baseline affective schedule, and the implicit responses of people to affective content. We infer the *big-five personality traits* [115] and the *Positive and Negative Affect Schedule (PANAS)* [179] by analyzing the features extracted from three modalities, namely (i) EEG signals, (ii) peripheral physiological signals (ECG, GSR), and (iii) facial landmark trajectories in response to 16 emotional videos.

We performed a mutual information (MI) analysis between different modalities and (i) arousal, (ii) valence, and (iii) dimensions of personality and PANAS. The analysis shows that the implicit responses are informative of the emotional state and the personality/PANAS of individuals. Therefore, the personality traits and PANAS could be revealed at the presence of emotions. We show that (i) emotional states have correlations with personality/PANAS and (ii) emotional state of individuals have normalized mutual information of about 0.5 with different modalities. Therefore, we expect similar levels of mutual information between modalities and personality/PANAS; this is shown to be true especially for peripheral physiological signals.

The main findings can be summarized as follows: *(i)* all three modalities have high mutual information with the dimensions of personality and PANAS, however the relation is not always linear. *(ii)* the peripheral physiological features have relatively higher mutual information with the dimensions of personality traits and PANAS than the other proposed features; *(iii)* due to strong linear relations between (a) EEG and openness and (b) peripheral physiological signals and extroversion, we achieved remarkably high mean F1-scores (about 70%) on the prediction of high/low extroversion/openness with a simple linear method.

The remainder of this chapter is structured as follows: section 3.2 summarizes previous research efforts in both (i) personality assessment methods and (ii) emotion measurement through psycho-physiological signals; section 3.3 provides an overview of the experimental protocol we followed; section 3.4 describes the data pre-processing and feature extraction steps taken followed by the the mutual information analysis; finally, after reporting experimental results in section 3.5, we discuss them along with the future research directions in section 3.6.

3.2 Related Works

In this section, we review the state of the art on measurement and prediction of affective behavior and personality.

3.2.1 Measuring Emotion

Emotions have a large impact on how we experience events in our life. Any behavior and environmental stimulus may have a psychological effect on us and may influence our interpretation of the environment and our consequent behavior. Knowing how people feel is helpful in improving interactions both in human-human and in human-computer interaction. Mainly two concepts are used in the literature of affect computing; One classifies emotions in six distinct universal groups [47]: happiness, sadness, anger, fear, surprise and disgust. The other is a dimensional model of emotion that is developed for the continuous measurement of affect. Wundt [185] proposed that emotions could be classified along three dimensions: pleasure, arousal and dominance. Bradley et al. [21] displayed emotions on a two-dimensional plane with the two axes valence (unpleasant-pleasant) and arousal (calm-aroused). Traditional methods to measure emotions are based on questionnaires. In order to detect emotion changes, it is useful to determine baseline levels of positive and negative affect the participants usually experience as it was done by Watson et al. [179] who developed the Positive and Negative Affect Schedule (PANAS).

To avoid the bias that can occur when people rate what they think they are supposed to feel instead of what they actually feel, emotions need to be decoded implicitly. Recent methods use physiological responses [89] or monitor users' facial expressions [80] since both (especially the former) are difficult to control.

Different affective states are correlated to changes in communicative signals such as speech, body language and facial expressions. An extensive review is given in [193]. Many researchers used the implicit responses acquired through psycho-physiological signals to predict the emotional states of humans [105, 90]. Lisetti and Nasoz [105] employed wearable devices to collect the physiological signals such as Galvanic Skin Response (GSR), heart rate (ECG), and skin temperature in order to predict basic emotions. They achieved a maximum 84% emotion recognition accuracy. Abadi et al. [2] measured emotions on the Arousal-Valence dimensions. They compiled a dataset with 30 subjects and used Magnetoencephalogram (MEG), Near Infra-Red (NIR) facial video, Electrooculogram (EOG), Electromyogram (EMG) and ECG responses for 36 emotional movie clips. Koelstra et al. [89] used EMG, EOG, blood volume pulse (BVP), skin temperature, and GSR to predict the emotional state of 32 participants upon watching on music videos. Soleymani et al. [149] created the MAHNOB-HCI multimodal database presenting emotional video clips to participants and collecting physiological signals to predict the emotional state. In this manuscript we take a step forward. By using a subset of videos from [2] and [149] we show that emotional states share high mutual information with personality and PANAS.

3.2.2 Personality Assessment

The Big-Five or the five-factor model describes human personality in terms of five dimensions: Extraversion (sociable vs. reserved), Agreeableness (compassionate vs. dispassionate and suspicious), Conscientiousness (dutiful vs. easy-going), Neuroticism or emotional stability (nervous vs. confident), and Openness or Creativity (curious vs. cautious) [115]. The traditional method to measure these personality dimensions has been through the use of questionnaires or self-reports. Other works used word frequencies in texts, non verbal communication aspects and body language cues for automatic personality recognition. There are few (if any) studies so far covering the connection between physiological signals and personality. The recent review [171] covers most of literature dealing with personality computing. Mairesse et al. [112] used acoustic and lexical features to develop classification, regression and ranking models for personality recognition. Social media are used as well to predict personality, especially with the increasing use of smart-phones that can be employed to measure different aspects of communication activities such as calls, instant messages and even frequency of speaking or proximity to other people in their social network. Srivastava et al. [152] presented a novel method for automating personality questionnaire completion utilizing behavioral cues for 50 movie characters, but this was not used in a real-life scenario.

Relationships between personality traits and user responses are mainly reported on Neuroticism and Extroversion [115, 85, 127]. Stenberg [154] reported relations between personality and arousal in an EEG-based study. According to [154], lower arousal levels are observed for extraverts as compared to introverts, while Neuroticism is associated with high arousal especially for negative valence stimuli. Gilbert [62] used active and passive coping tasks as stimuli and found that heart rate and skin conductance correlate with Extroversion and Neuroticism. Stough et al. [155] found correlations between Openness and Conscientiousness with EEG signals when using photic driving. While previous studies mainly concentrated on finding correlations between implicit responses and personality we employ the implicit responses for the prediction of personality traits. Additionally, to the best of our knowledge, we are the first to use pycho-pysiological responses to predict PANAS.

3.3 Experimental Protocol and Rating Analysis

3.3.1 Used stimuli and experimental protocol

Selected stimuli

Our objective for stimuli selection was to select videos that covered well the arousal and valence (AV) space. For each quadrant of the AV space (High Arousal-High Valence (HAHV), Low Arousal-High Valence (LAHV), Low Arousal-Low Valence (LALV), and High Arousal-Low Valence (HALV)) 3 videos were selected from the 36 videos used in [2]. This selection was made based on the self-assessment of 80 participants. Additionally, one video for each quadrant was selected from the ones used in MAHNOB-HCI [149], giving a set of 16 videos (4 per each quadrant). Selected videos (51s-150s long ($\mu = 86.7$, $\sigma = 27.8$)) are listed in Table 3.1. Each video is given an ID that is used to refer to it in the remainder of the chapter.

Materials and Setup

Experiments were performed in a laboratory environment. Physiological signals were obtained using wearable sensors. EEG was recorded using an Emotiv EPOC Neuroheadset[1] (14 channel {AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4}, 128 Hz, 14 bit resolution). For ECG and GSR signals recording, two extended Shimmer 2R[2] platforms (12 bit resolution) working at 256 and 128 Hz, were used. A MATLAB[3] based platform running on a PC (Intel Core i7, 3.4 GHz) was used to (i) present the stimuli, (ii) obtain and synchronize the signals, and (iii) get the users' ratings. Subjects were seated approximately at 2 meters from the screen (40-inch,1280 x 1024, 60 Hz) where stimuli were presented at the maximum scale that conserved the original aspect ratio. The sound volume was adjusted for each participant to a comfortable level. Frontal face video was recorded with a JVC GY-HM150E camera.

Experimental protocol

35 healthy participants (12 female), aged between 24 and 40 (mean age 28.85), participated in the experiment.

Preparation: Each participant was informed of the experimental protocol and signed a consent form before she/he was led into the experiment room. The experimenter explained the scales used and how to fill the self-assessment form. Then the sensors were placed and their signals checked. The participant started

[1]http://www.emotiv.com/

[2]http://www.shimmersensing.com/

[3]http://www.mathworks.co.uk/

Table 3.1: The Video Clips Listed with Their Sources (Video IDs are stated in parentheses). In the category column, H, L, A, and V stand for high, low, arousal and valence respectively.

Category	Excerpt's source
HAHV	Airplane (1), When Harry Met Sally (2), Hot Shots (3), Love Actually (4)
LAHV	August Rush (5), Love Actually (6), House of Flying Daggers (7), Mr Beans' Holiday (8)
LALV	Gandhi (9), My girl (10), My Bodyguard (11), The Thin Red Line (12)
HALV	Silent Hill (13), Prestige (14), Pink Flamingos (15), Black Swan (16)

the experiment once the experimenter left the room.

Experiment pipeline: The recording session started with an initial emotion self-assessment. The 16 videos were presented in a random order in trials consisting of a 5 second baseline recording (fixation cross), the presentation of a short video (see 3.3.1), followed by the video emotion self-assessment.

Participant self-assessment

At the beginning of the experiment and at the end of each trial, participants performed a self-assessment of their affective state. Self-assessment manikins (SAM) [21] with continuous sliders at the bottom were used to visualize the scales of arousal, valence and dominance. Participants moved the sliders to specify their self-assessment level in a continuous scale. Arousal ranges from *very calm: 1* to *very excited: 9*, valence from *very negative: 1* to *very positive: 9*, and dominance from *overwhelmed with emotions: 1* to *in full control of emotions: 9*. In addition, each participant was also asked to select one or more emotional keywords (Neutral, Disgust, Happiness, Surprise, Anger, Fear, and Sadness) they considered that described their emotional state (1: if chosen, 0: otherwise). The whole experiment including the preparation steps took 50 minutes on average per person.

3.3.2 Personality and PANAS evaluation

Big Five Personality

The Big Five personality traits were measured using the big-five marker scale (BFMS) questionnaire [133]. For each personality trait ten descriptive adjectives were rated on a 7-point Likert scale and the mean was calculated. The distributions of personality measures over all participants are presented in Figure 3.1(a). While for Extroversion and Emotional Stability they are more equally distributed, the average scores for Agreeableness, Conscientiousness and Creativity are more clustered with a higher average close to 5.

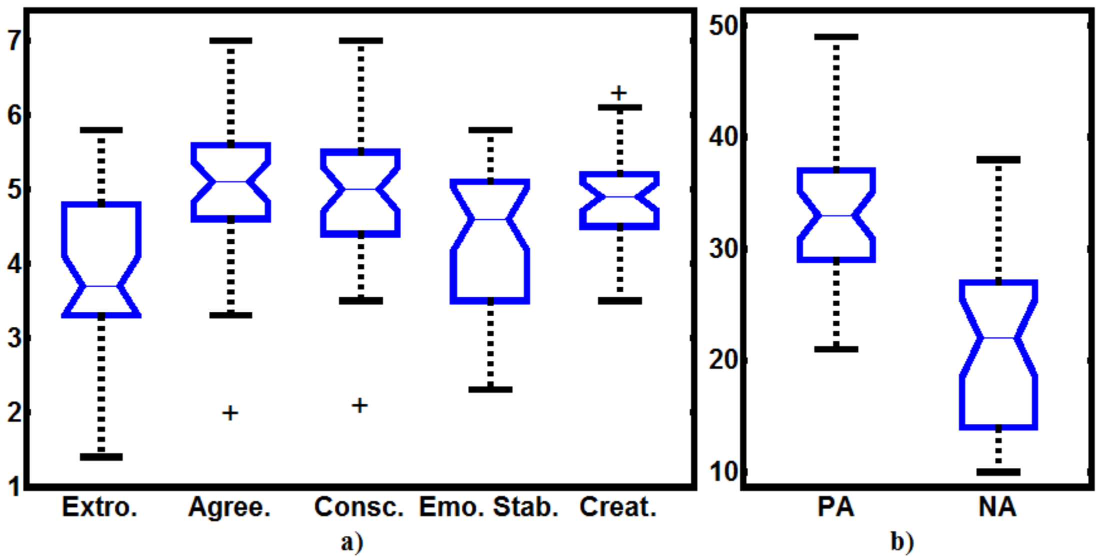


Figure 3.1: a) Distribution of the big-five personality traits. b) Distribution of the average Positive and Negative Affect (PA and NA).

PANAS

We used the General PANAS questionnaire [179] consisting of 10 questions each to access the positive and the negative affect. The participants filled an online form rating their general feelings on a 5-point intensity scale using questions like ”Do you feel in general...?”. The positive feelings asked are: active, alert, attentive, determined, enthusiastic, excited, inspired, interested, proud, strong. The negative ones asked are: afraid, scared, nervous, jittery, irritable, hostile, guilty, ashamed, upset, distressed. The resulting positive and negative affect measures are mostly independent as shown in [178]. This allows to investigate both aspects independently. The correlation coefficient is 0.12 which is similar to the ones reported in the literature [178]. PANAS is calculated by summing the values (between 1 and 5) of all 10 questions for PA and NA respectively resulting in values between 10 and 50. The distribution and average for PA and NA is consistent with the literature as well [178]. The mean PA is 32.9 while the mean NA is lower (21.3) as presented in Figure 3.1(b).

3.3.3 Affective Rating Analysis

We evaluated the suitability of the presented stimuli in terms of their power to evoke emotions in participants. The mean and standard deviation of participants' self-assessments of arousal and valence for each video is reported in Table 3.2. Upon calculating the mean for emotional keywords of each video over participants, the mean values were normalized to sum up to 100 to get the percentage of reported emotional keywords (see Table 3.2). According to table 3.2, the chosen stimuli for the four quadrants of the AV space (LALV, HALV, LAHV, and HAHV) generally resulted in the elicitation of the target emotions, and the four quadrants are covered. The relatively lower values of dominance self-assessments over HALV suggest that the participants were more emotionally touched by negative videos.

Among the emotional keywords (adapted from [47]), *happiness* is the only *positive* keyword. We observe *happiness* to be among the dominant emotional keywords chosen for all the HV videos. All HA videos are associated with *surprise* (often as the second dominant keyword), as surprise is characteristic to *excitement*. Interestingly, all the LA videos are labeled with the *neutral* keyword (often as the second dominant keyword), which is due to the lower intensity of

Table 3.2: The mean and standard deviation of participants ratings (range = [1,9]), over arousal and valence dimension for each video is reported. Moreover, the table includes the normalized histogram of the selected emotional keywords (**N**eutral, **A**nger, **D**isgust, **F**ear, **H**appiness, **S**adness, and Sur**P**rise) for each video clip. The dominant emotional keywords of each video are bolded.

	Video ID	**Arousal** $\mu \pm \sigma$	**Valence** $\mu \pm \sigma$	**Dominance** $\mu \pm \sigma$	% **N**	% **A**	% **D**	% **F**	% **H**	% **S**	% **P**
HAHV	1	5.5 ± 1.8	5.8 ± 2.0	5.4 ± 2.3	18	8	3	6	**31**	2	**32**
	2	6.0 ± 1.5	6.8 ± 1.3	4.7 ± 1.9	14	1	0	2	**49**	0	**34**
	3	5.5 ± 1.7	6.4 ± 1.3	5.3 ± 2.1	14	2	2	3	**47**	0	**32**
	4	5.4 ± 1.5	7.7 ± 1.0	5.0 ± 1.8	15	0	0	0	**60**	0	**25**
LAHV	5	3.8 ± 1.7	7.0 ± 1.2	6.2 ± 2.2	**44**	0	0	0	**52**	0	4
	6	4.1 ± 1.8	7.7 ± 1.0	5.4 ± 2.1	**32**	0	0	0	**54**	7	7
	7	3.7 ± 1.5	7.3 ± 1.0	5.8 ± 2.0	**27**	0	0	2	**57**	8	8
	8	4.4 ± 1.6	7.0 ± 1.0	6.1 ± 1.9	**33**	0	0	0	**61**	0	6
LALV	9	4.4 ± 1.6	3.9 ± 1.5	6.0 ± 1.8	**32**	**26**	8	5	3	13	13
	10	5.2 ± 1.6	3.5 ± 1.3	5.0 ± 2.1	**24**	7	2	2	0	**63**	2
	11	5.0 ± 1.5	3.4 ± 1.4	5.3 ± 2.2	**27**	**33**	9	5	3	12	11
	12	4.2 ± 1.4	3.6 ± 1.1	5.5 ± 1.9	**26**	9	5	3	0	**55**	2
HALV	13	6.8 ± 1.4	3.4 ± 1.8	4.5 ± 2.0	17	3	13	**38**	0	1	**28**
	14	5.9 ± 1.5	3.2 ± 1.4	4.7 ± 2.0	11	9	4	20	0	**33**	**23**
	15	5.5 ± 1.4	2.9 ± 2.1	4.0 ± 2.0	7	10	**58**	0	6	2	**17**
	16	6.5 ± 1.7	3.1 ± 1.4	4.4 ± 1.8	8	4	14	**41**	1	7	**25**

emotion in *lower arousing* videos [175].

In LV videos, the underlying negative emotional keywords (sadness/disgust/fear/anger) are often the most dominant reported ones. We observe that the *anger* keyword is only dominant in LA videos. that indicates that the 9th and the 11th video clips involve low intensities of *anger* and evoke *irritation/pity* more than *rage/anger*.

T Wilcoxon signed-rank tests showed that low and high arousal stimuli induced different valence ratings ($p < .005$ and $p < .001$). Similarly, low and high valence stimuli induced different arousal ratings ($p < .0001$ and $p < .0001$).

Table 3.3: Observed significant correlations between Personality/PANAS dimensions versus explicit emotional responses (self assessments). Each reported item starts with a letter indicating the emotion dimension (A,V, and D for arousal, valence and dominance, respectively), followed by the ID of the video for which the correlation is observed (in parentheses the correlation value is stated)

Dimension	**Observed significant** ($p < 0.05$) **correlations**
Extroversion	V16 (0.19) D6 (-0.19)
Agreeableness	A2 (-0.08) A8 (-0.42) V3 (-0.04) D16 (-0.04)
Conscientiousness	A7 (0.06) V16 (-0.06)
Emotional Stability	A1 (0.41) A12 (0.13) A14 (0.28) D1 (-0.34)
Creativity (openness)	V6 (0.27) V11 (0.16)
Positive Affect Schedule	A3 (-0.08) A6 (-0.26) V16 (0.04)
Negative Affect Schedule	A15 (-0.22)

The distribution of the individual ratings per conditions shows a large variance within conditions. This can be explained by between-stimulus and between-participant variations. We investigated the mean inter-correlation of the arousal and valence scales over participants. The mean of the subject-wise inter-correlations between the scales is -0.168. The correlation is significant ($p < .05$) - this is consistent with other studies [89]. Even though the arousal and valence scales are not independent, their negative correlation is quite small implying that participants could differentiate between them.

We measured the Spearman's correlation between the affective ratings of each video provided by the 35 participants versus the personality traits as well as PANAS measures over the 35 participants. The significant observed correlations ($p < .05$) are reported in Table 3.3. Previous works established a link between psycho-physiological signals and affective states [125, 87, 89, 149]. Therefore, the obtained correlations between explicit emotional responses (affective self-assessments) and the personality and PANAS dimensions reported in Table 3.3, suggest that the implicit emotional responses (i.e. psycho-physiological signals) should also relate to personality and PANAS dimensions. In the next sections, we present a method to predict the personality and PANAS dimensions using a person's implicit responses to emotional videos.

3.4 Data Analysis

We used 3 modalities to record the implicit emotional responses of people: (i) EEG, (ii) peripheral physiological signals (ECG and GSR), and (iii) facial videos. We extracted state of the art *affective features* from different modalities for our analysis. In this section we first describe in detail the extracted features from the employed modalities and then analyze their mutual information with the different affect/personality/PANAS dimensions. To avoid any bias due to different video lengths, all the features are calculated using the responses over the last 50 seconds of the videos.

3.4.1 EEG Signal Processing

EEG measures the electrical activity on the scalp. For obtaining features from the EEG signals, the EEG data was processed similarly to [89], using the sampling frequency of 128 Hz. To correct for stimulus-unrelated variations in power over time, the EEG signal from the five seconds before each video was extracted as baseline. Using the Welch method with windows of 128 samples, the frequency power of trials and baseline signals between 3 and 47 Hz was calculated. The baseline power was then subtracted from the trial power, yielding the change of power relative to the pre-stimulus period. These changes of power were averaged over the frequency bands of theta (3-7 Hz), alpha (14-29 Hz), beta (8-13 Hz), and gamma (30-47 Hz). Additionally, the spectral power asymmetry between 7 pairs of electrodes in the four bands was calculated. The complete set of features is listed in Table 3.4.

3.4.2 Peripheral Physiological Signal Processing

We used the methods reported by Kim and Andrè [87] to preprocess the ECG and GSR signals and then extract the features.
Galvanic Skin Response: GSR provides a measure of the electrical resistance of the skin. This resistance varies due to changes in perspiration that are controlled by the sympathetic nervous system (SNS). The changes in GSR are related to the presence of emotions such as stress or surprise while the mean of the GSR signal is related to the level of arousal [95]. In our setup the electrical resistance between two electrodes positioned on the middle phalanges of the middle and index fingers is measured as the GSR signal.

Following [87] we calculated the skin conductance (SC) from GSR and then normalized the SC signal. We low-pass filtered the normalized signal with 0.2 HZ and 0.08 Hz cut-off frequencies to get the low pass (LP) and very low pass (VLP) signals, respectively. Then, we detrended the filtered signals by removing the continuous piecewise linear trend in the two signals. We calculated 31 GSR features employed in [89, 149] and that are listed in Table 3.4.
Electrocardiogram: The ECG signal was recorded using three electrodes attached to the participant's body. Two of them were placed on the right and left arm crooks and the third one was placed to the left foot as reference. This setup allows precise identification of heart beats. Using the method reported in [87] we accurately localized the heart beats in ECG signals (R-peaks) to calculate the inter beat intervals (IBI). Using IBI values, we calculated the heart rate (HR) and heart rate variability (HRV) time series. Following [149, 87] we extracted 77 features listed in Table 3.4. In this study we use the concatenation of ECG and GSR features as the peripheral physiological features.

Table 3.4: Extracted affective features for each modality (feature dimension stated in parenthesis). Computed statistics are: mean, standard deviation (std), skewness, kurtosis of the raw feature over time, and % of times the feature value is above/below mean ± std.

Modality	Extracted features
ECG (77)	root mean square of the mean squared of IBIs, mean IBI, 60 spectral power in the bands from [0-6] Hz component of the ECG signal, low frequency [0.01,0.08]Hz, medium frequency [0.08,0.15], and hight frequency [0.15,0.5] Hz components of HRV spectral power, HR and HRV statistics.
GSR (31)	Mean skin resistance and mean of derivative, mean differential for negative values only (mean decrease rate during decay time), proportion of negative derivative samples, number of local minima in the GSR signal, average rising time of the GSR signal, spectral power in the [0-2.4] Hz band, zero crossing rate of skin conductance slow response (SCSR) [0-0.2] Hz, zero crossing rate of skin conductance very slow response (SCVSR) [0-0.08] Hz, mean SCSR and SCVSR peak magnitude
EEG (84)	4 bands (theta, alpha, beta, and gamma) spectral power for each electrode. The spectral power asymmetry between 7 pairs of electrodes in the four bands.
Facial Landmark tracks (72)	Statistics concerning horizontal and vertical movement of 12 motion units (MUs) specified in [80].

3.4.3 Facial Video Analysis

We used state of the art methods to initialize and track the facial landmarks and then we extracted statistic measures over 12 motion units (MU) as facial features.

Facial landmark tracking: We extracted the time series of facial landmark location tracks. Before applying the tracking methods, we used the Robust Cascaded Pose Regression (RCPR) [29] with detection model from [188] and the SDM [186] face alignment methods over the first few frames of the facial video. Both of the methods detect the facial landmarks and work in a cascaded way.

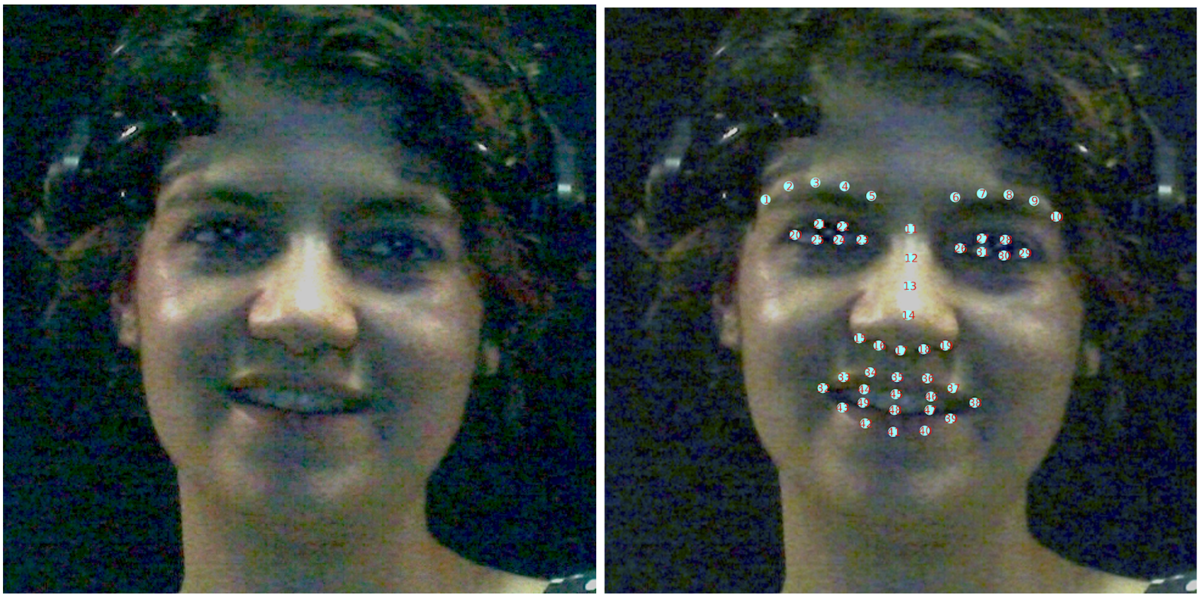


Figure 3.2: Left Image: A sample frame of a participant's facial video. Right Image: The output of the SDM facial landmark detection algorithm. The ID of the location of the 49 landmarks are visible under zoom.

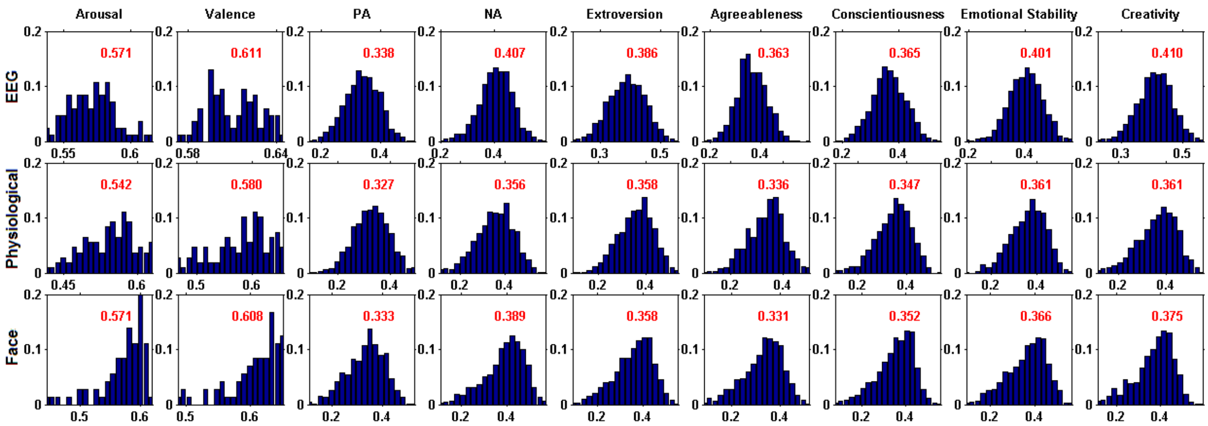


Figure 3.3: The normalized histograms of normalized mutual information between each modality and affect/PANAS/persoanlity dimensions. The first moment of each distribution is shown in red (best viewed under zoom).

SDM uses only a shape inside the face bounding box as initialization of the face shape (locations of the facial landmarks). In each cascade, based on the calculated Histograms of Oriented Gradient (HoG) features [42] that are calculated. In the surrounding of each landmark, a linear regression is applied. *RCPR* uses several face shape initializations, normalized by the face bounding box. At each cascade, random ferns are used as the primitive regressor for calculating the update. Upon extracting the landmarks using the SDM and RCPR, we validate the correctness by calculating the difference of the locations of their common landmarks. When the difference is smaller than a threshold (set empirically), we use the SDM method to obtain the tracks. Otherwise, the landmarks are set manually. In our experiments, only a few videos, mainly ones in poor lighting conditions needed to be manually checked. The SDM outputs the track of 49 inner facial landmarks using the pixel locations as reference. The landmark detection sample over a frame of a participant's facial video is shown in Figure 3.2.

Processing the facial landmark tracks: To discard the head movement artifact from the facial landmark tracks, we subtracted the track of the nasion (landmark #11) from all the other tracks. Then each track was low-pass filtered with a cut-off frequency of 1Hz. The tracks are used to determine the time series of 12 motion units (MUs) according to [145, 80]. Statistics over the 12 time-series are used as features (see Table 3.4).

3.4.4 Mutual Information Analysis

We performed a mutual information analysis between the extracted features from the three modalities versus affect/PANAS/personality dimensions (9 in total). Mutual information (MI) between two random variables measures how much information is known about one of the random variables when the other is known. The function that defines the MI of two random vectors x and y is defined by:

$MI(x,y) = \sum_{i,j} p(x_i,y_j) log \frac{p(x_i,y_j)}{p(x_i).p(y_j)}$ where $p(x,y)$ is the joint probability distribution and $p(x)$ and $p(y)$ are the respective marginal probabilities. After calculating the MI between each modality and the affect/PANAS/personality dimmensions, we calculate the normalized mutual information (NMI) index [156] using the following equation: $NMI(x,y) = \frac{MI(x,y)}{\sqrt{H(x)H(y)}}$ where $H(x)$ and $H(y)$ are the entropies of x and y. We used the MIToolbox [23] to calculate the MI index and entropy values after normalizing x and y to $[0,10]$. Figure 3.3 presents the normalized 20-bins histograms of the distribution of NMIs for every modality and dimension. The histogram normalization allows a better comparison given the different number of features for every modality. For each normalized histogram, we also calculate the first moment (indicated in red text in Figure 3.3), to summarize the distribution of NMIs. The presented results in Figure 3.3 suggest that (i) the extracted features from different modalities share information with arousal and valence dimensions and hence they contain information about affective state of the participants. The results also suggest that (ii) the features contain information about the participants' personality/PANAS measures. From the two observations we may expect to obtain above-chance prediction of personality and PANAS dimensions, and that may be with the help of affective information included in the extracted features.

3.5 Experimental Setup and Results

In this section we describe our method for the prediction of personality/ PANAS based on the extracted features in a *leave-one-subject-out* cross-validation schema.

3.5.1 Personality/PANAS recognition

Each participant watched 16 emotional video clips and for each participant we have five measures for the big-five personality traits and two measures for PANAS. To this end we extracted the features listed in Table 3.4 of 35 participants for each of the 16 emotional videos.

Recognition tasks: We associate all the emotional responses of a participant to his/her personality/PANAS measures and we propose a method that can predict the measures of a new test participant based on the available training data.

In total we have seven recognition tasks; five for personality traits and two for PANAS.

Experimental Schema: We use a leave-one-subject-out cross-validation schema to validate our proposed method for solving the *recognition tasks*. Assuming the dataset includes N participants as our samples, in each iteration of the cross validation, we take out one sample as the test sample and use the rest as training samples. We train a linear regression model using the N-1 training samples and we predict the measure for the test sample. After completing all the N iterations we dichotomize the prediction and the ground truth values using the median criteria as threshold to divide the samples into high/low classes (e.g. high/low score on extroversion). We then use mean F1-score of high/low classes to evaluate the quality of the predictions. To more reliably report the performance of our method, we ran the whole cross-validation process 1000 times. In each run, 31 subjects were randomly chosen as samples ($N = 31$) from the 35 available participants. In Table 3.5 the mean and standard deviation over the obtained results from 1000 runs is presented. The table also includes the random baseline results that are obtained using three methods for the sake of comparison; (i) random voting, (ii) majority class voting, (iii) class distribution voting according to [89]. We also employed a t-test to probe which of the results has a distribution with a significantly ($p < 0.001$) higher than chance level (0.50) mean. The distributions for which the lower bound of the confidence-intervals are more than 0.55 are bold.

Method in Detail: For a certain recognition task (e.g. recognition of extroversion) and a certain modality (e.g. EEG), all the 16 feature vectors in response to 16 emotional videos are taken as samples of a participant. The 16 samples of a participant are associated with the measure of the target dimension (e.g. extroversion). For each participant, the features extracted from each modality in response to 16 video clips are mapped to the range of $[-1, 1]$ over the 16 clips. The normalization removes the subjective artifacts and puts the focus of the pattern recognition on differentiating between the responses to different affective videos. During the training, after pooling all the samples ($30 \times 16 = 480$ in total) from the training subjects, we calculate the $z-score$ of features along all the samples. The same parameters of the second normalizations (μ and σ) are used to map the samples of the test subject. We also normalized the scores associated to train/test samples with the parameters of the map to $[-1, 1]$. Then we used SVD decomposition to solve the following equation for W_{Tr} (the regression weights):

$$W_{Tr} \times [\mathbf{1}\ D_{Tr}] = S_{Tr} \tag{3.1}$$

where D_{Tr} contains the normalized training samples in its columns and S_{Tr} contains the normalized target dimension scores of the train samples in one row vector. We use W_{Tr} to predict P_{Ts}, the prediction of the target dimension of the test subject, using the following equation:

$$W_{Tr} \times [\mathbf{1}\ D_{Ts}] = P_{Ts} \tag{3.2}$$

where D_{Ts} includes the normalized test samples in its columns. Table 3.3 suggests that the responses to some videos are more useful for the prediction of the target dimension. Therefore, we select a set V of 3 videos that yield the best performance over training samples. Then, we calculate the *median* of the predictions for the videos in V as the estimation of the score for the target dimension of the test subject.

3.5.2 Discussion on the Results

Our method for the prediction of different personality/PANAS dimensions is based on a linear regression, therefore it is computationally very cheap but cannot capture nonlinear relations. We observed that different modalities share information (Figure 3.3) with the personality/PANAS dimensions. However not all of the relations are linear. The obtained results presented in Table 3.5 suggest that the extracted features from peripheral physiological signals have

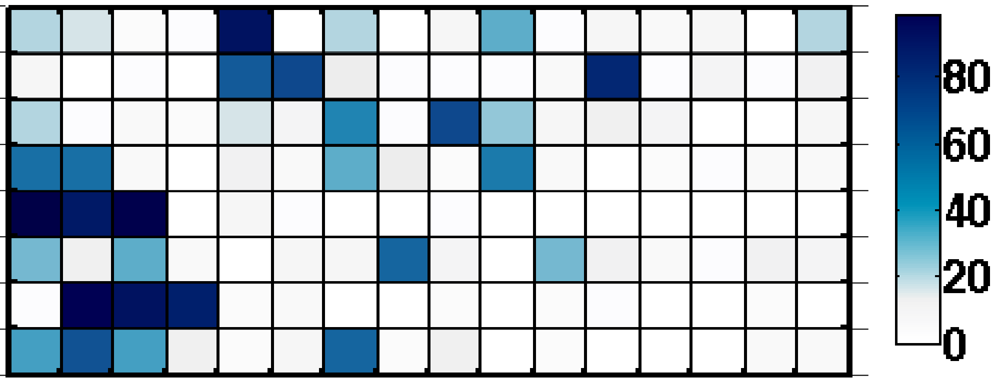
Table 3.5: Mean and standard deviation over 1000 independent runs. In each run the performance of a leave-one-subject-out cross validation using 31 participants out of 35 people is measured. The mean-F1 scores of binary classes are used to evaluate the performance. The results of random prediction baseline using three methods; random voting, majority class voting and class ratio voting are also reported.

Modality	**Ext.**	**Agr.**	**Con.**	**Emo.**	**Cre.**	**PA.**	**NA.**
Emotive EEG	0.44± 0.07	**0.60± 0.07**	0.53± 0.08	0.53± 0.07	**0.69± 0.06**	0.38± 0.07	0.49± 0.06
Physiological signals	**0.70± 0.08**	0.50± 0.08	0.53± 0.09	**0.58± 0.08**	0.53± 0.08	**0.60± 0.07**	**0.58± 0.09**
Facial Tracks	0.50± 0.07	**0.58± 0.06**	0.38± 0.09	0.45± 0.07	0.52± 0.08	**0.59± 0.08**	0.48± 0.09
Random Baseline	0.49± 0.09	0.50± 0.09	0.50± 0.09	0.49± 0.09	0.49± 0.09	0.50± 0.09	0.49± 0.09
Majority Baseline	0.50± 0.01	0.50± 0.01	0.50± 0.01	0.50± 0.01	0.50± 0.01	0.50± 0.01	0.50± 0.01
Class Ratio Baseline	0.34± 0.01	0.35± 0.01	0.36± 0.02	0.35± 0.01	0.36± 0.01	0.35± 0.01	0.36± 0.01

more (strong) linear relation with different dimensions, particularly with extroversion scores. Spectral power features extracted from EEG responses seem to have strong linear relation with openness. This result is in line with the related exploratory studies [155, 62, 154]. Particularly, (i) Stough et al. [155] found correlations between EEG signals and openness and (ii) Gilbert [62] found that heart rate and skin conductance correlate with Extroversion and Neuroticism. As future work, we will investigate capturing the nonlinear relations between different modalities and target dimension.

Table 3.6: Importance of the role of the 16 videos for the best prediction performance reported in Table 3.5. The values are presented in terms of colors from 0 (white) to 100 (dark blue). The values indicate the mean percentage of times that a video was selected for the prediction of test samples of a certain dimension, while using the features from a certain modality. The reported results are the mean percentage over 1000 independent runs.

		HAHV				LAHV				LALV				HALV			
Dim.	Modal.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Ext.	Phys.																
Agr.	EEG																
Agr.	Face																
Emo.	Phys.																
Cre.	EEG																
PA.	Phys.																
PA.	Face																
NA.	Phys.																



3.5.3 Chosen videos during the predictions

As mentioned above, each prediction (over a dimension) is based on the predictions over a test person's responses (collected through a modality) to 3 videos, (set V) out of the 16 presented videos. The chosen three videos are the ones that best help the prediction of the dimension. It is interesting to know which videos were *usually* selected for the successful predictions over a dimension. Over the 1000 runs for the prediction of a dimension using a modality, we counted the occurrence of all videos in the chosen set V. Then the percentage of *the times that each video is chosen for the prediction* is calculated and reported in Table 3.6. Since each prediction involves 3 videos, the sum over the entries in each row of Table 3.6 is equal to 300%.

Discussion: For distinguishing between *extroverts/introverts*, videos from all categories were involved. However, low arousal (LA) videos were chosen more often and particularly *August rush (happy)* and *My girl (sad)* were the most effective videos for the prediction of *extroversion*.

For the prediction of *agreeableness*, mainly low arousal videos were selected. The difference between the chosen videos for different modalities (EEG and Facial landmark tracks) for the prediction of agreeableness suggests the presence of complementary information in different modalities and encourages the fusion of information for future extensions of this study.

The important videos in estimating the level of *emotional stability* are selected from all the four quadrants but HALV. HALV videos in our study were rarely chosen for the predictions. The reason may be that the negative videos in HALV (scary/disgusting/stressful videos) are very emotional that touch the majority of population similarly and hence, the responses of different people to HALV videos are not very useful for the predictions. A support for the last statement is that the HALV videos are shown in Table 3.2 to be the most emotional ones (with lower values of dominance).

Interestingly, the top videos for the prediction of *openness* are only from HAHV (funny) videos. The observation suggests that the reaction to funny moments in videos is very useful for the prediction of *creativity*.

Positive affect (PA) and negative affect (NA) schedules were mainly estimated through (physiological) responses to positive videos. However in the estimation of the level of PA through facial tracks, *funny videos* (HAHV) had the main role, suggesting that facial expressions to humorous stimuli are distinctive for general positive affect.

3.6 Conclusion

This study proposes a method for predicting users' big-five personality traits and PANAS of people based on the analysis of their implicit responses to emotional videos. We used 16 emotional videos to evoke emotions in people and recorded the implicit responses through wearable EEG, GSR, and ECG sensors, as well as facial videos. We observed that all the employed modalities share high information with the personality and PANAS dimensions and we showed in some cases that a linear model can model well the relation. We tried to capture the linear relations with a linear regressor to predict the correspondent dimensions. The accurate prediction of personality traits and PANAS can later be used (i) to profile people in human-computer interaction and (ii) to develop cross-subject personality/PANAS predictors. Even though we could already show mutual information among constant characteristics (personality traits and General Affect) with changing reactions (EEG, physiological signals and facial expressions), we believe that by using nonlinear regression methods we can obtain even better results. This will be addressed in future work to contribute to better user profiling in human computer interactions.

Chapter 4

Respiration as indicator of physiological states

Stress is a recurrent factor in modern life that people try to reduce. To monitor stress we propose to use physiological signals such as breathing patterns. The breath rate depends on the psychological state of the person, but in contrast to other physiological signals it can also be actively influenced. For example, deliberate slow breathing can be a calming process relaxing body and mind. We investigated the benefits of improving breathing patterns at Spire in San Francisco. We conducted a study investigating how to increase the awareness of breathing patterns and consequent effects on reducing stress in a real life setting. We designed a wearable device, the Spire sensor, which detects breathing patterns and a respective application that provides correspondent information and suggestions to improve breathing behavior. To assure that no additional stress is imposed on the user through the interface, we considered corresponding design principles to avoid additional mental effort for the users.

4.1 Breathing and its connection to stress

As we saw in the first part of the thesis, physiological data contain information about emotion and personality (Chapters 2 and 3). One important emotional state, which is often present in today's population, is stress, often caused by frustration or cognitive overload. In this chapter we analyze how to use technology in order to reduce stress by using the information of a persons physiological states measured via breathing patterns.

So far, many physiological effects have been correlated to emotions [105], while respiration data have been studied less intensively. Breathing behavior changes with emotions, especially with the stress level someone experiences [37]. Stress is related to diseases (burn out, sleeping disorders or even heart attacks) because permanent negative emotions can cause negative physiological and psychological effects [61]. Therefore, researchers have searched for easy strategies to reduce stress [121]. In this context, two aspects of breathing are important: (i) breathing patterns are correlated to the stress level [37], however (ii) breathing patterns can also be actively changed because they are under conscious control. Consequently, breathing cannot only be used to understand the current mental state and stress level, but also has the potential to change people's physiological and psychological states by relaxing the body, e.g. as practiced in yoga [25]. By increasing awareness and learning to identify bad breathing habits (breath-holding, hyperventilation, etc.) we aim at helping to return to natural, effortless, and deep breathing, reducing negative physiological effects such as fatigue and negative mood and their consequences [101].

We conducted a user study to investigate the positive effects that controlled breathing provides in a stressful working environment. The study and its results are described in the following. Our preliminary results suggest that indeed breathing patterns are correlated to stress as previously identified by [25, 26].

4.2 Controlled Breathing and Physiological Effects

Physiological data are a good indicator of emotions and stress [148]. Respiration patterns are particularly interesting signals because they are easier to influence than other autonomic body responses as breathing is under cognitive control. Respiratory rate changes with body conditions and is used to monitor both illness and general wellbeing [4]. Stressed persons often breathe fast and irregular. Hence, increasing awareness of breathing can help to stabilize the breathing rate [37].

Reducing the breath rate actively leads to activation of the parasympathetic nervous system. It relaxes the body, therefore regulating the breath is a common practice to alleviate (physiological) stress symptoms [105]. Clark [36] showed in a clinical trial that deliberately breathing with a low breath rate causes even more reduction of anxiety and thereby helps to relax. Breathing techniques have been shown to have medical benefits, e.g. calm breathing is an effective non-pharmacological method to reduce blood pressure. Ten [69] or fifteen [117] minutes of conscious breathing per day to a calm paced rhythm proved to be beneficial for patients with hypertension. Ley [102] discusses the possibilities to condition changes in breathing behavior and found that indeed changes in aspects of breathing can be conditioned, for example to decrease the breathing rate. Therefore, training slower paced breathing should help reducing physiological arousal. In order to benefit from calm breathing also during performing other tasks, Moraveji et al. [122] aimed at integrating calm breathing into normal work processes. They integrated a visual feedback in the operating system of a computer to stimulate peripheral paced respiration. While following the visual breathing feedback, and breathing more calmly accordingly, the users were still able to work on other tasks.

It remains to investigate in more depth how respiration patterns are connected to emotions. Breathing more calmly has reportedly positive effects [36] but normally requires active attention [37]. Consequently, finding a way to monitor and raise awareness for breathing could be beneficial to reduce stress. Moraveji, the author of [122], therefore decided to co-found the start-up Spire[1] that the

[1]www.spire.io

author visited during a three month internship. We developed a breathing sensor to reduce stress by increasing awareness of breathing patterns and supporting users to change them actively, e.g. breathing more calmly.

4.3 Study

Nowadays people in many jobs are exposed to stressful working environments. Especially when working on a computer, they have to deal with multitasking and frequent distractions. The permanent pressure to be contactable at any time (i) increases distraction, (ii) decreases productivity and (iii) leads to a higher stress level. Regulating breathing can be beneficial for reducing stress [37], but acquiring new breathing habits and breathing in a more controlled way can be a challenging. Raising awareness about ones breathing behavior and correspondent feedback can therefore be of great help to improve breathing behavior leading to further benefits.

4.3.1 Hypotheses

Together with Spire we conducted a user study to test a new approach. By using the prototype of their wearable sensor, we measured the breath rate and provided feedback to the participants, creating awareness and incentivizing to breathe more frequently in a calmer way.

Our app is supposed to help people being more aware of their breathing habits and change them. The goal is to breathe more calmly, at least 20 minutes per day. Each half-minute that the participant was breathing with a low breath rate was counted. Consequently our research hypotheses were the following:

1. Using Spire, i.e. the sensor with the corresponding app, increases calm breathing and therefore the average breath is reduced.

2. Using Spire has positive psychological effects such as (i) increasing users' awareness of their breathing patterns and (ii) reducing their stress level.

4.3.2 Participant selection

We informed the employees of the company LinkedIn about our study and all people interested in participating were invited to fill in a selection survey. In order to be eligible, people had to own an iPhone 4S or newer, be at the office during the study period and be willing to fulfill all the tasks required by the experiment. 20 participants were selected (10 female, 10 male) but only 16 people completed the study without more than one missing day (out of 5 days). Their age ranged from 23 to 41 with an average of 31.

4.3.3 Sensor

To detect the breath rate we used Spire's prototype of a new breathing sensor; the newly developed wearable sensor Spire is light, small and non-intrusive. The sensor has a clip that is attached to the waistband of the trousers detecting abdominal movement that is caused by breathing, without needing any skin contact (Figure 4.1). It contains a pressure sensor and an accelerometer that collects data with a sample rate of 30Hz. By using this data one can calculate the breath rate. The accelerometer provides additional data about general movement and the amount of steps walked. With this data it is possible to detect whether a person was active (moving) or at rest (sitting, lying) during a specific time. The sensor connects via Bluetooth to the iPhone for which we designed an app that is supposed to help people being more aware of their breathing habits and change them.

4.3.4 App

The app consists of a single screen. In the center the number of minutes breathed calmly during the day is displayed and next to it the goal for that day. In the top a small animation of the actual breathing state is demonstrated and on the bottom tips and information are presented. According to their status, participants

(a) (b)

Figure 4.1: (a) Prototype Spire sensor with the clip to connect to the waistband. (b) Sensor attached to trousers.

received push notifications informing them when the Bluetooth connection was lost, the sensor was not placed correctly, they reached their daily goal or they had a long streak of calm breathing. If they did not collect any minutes of calm breathing for more than 1.5 hours they also received a notification.

4.3.5 Procedure

On the first day all participants gathered for an introduction. Each person received a Spire sensor and installed the Spire app on the phone. They learned how to use them and what other tasks they had to perform during the study. For one week they wore the sensor throughout the workdays and were advised to breathe deeply and calmly for at least 20 minutes per day.

In the evening participants received a daily report informing them about the amount of minutes they were breathing calmly and how they compared to the other participants. A diagram was shown that gave an overview of the average breath rate throughout the day. Additionally they had to fill in a questionnaire asking how they felt during the day and what impact the sensor had on their daily life.

4.3.6 Data collection

As mentioned before, the sensor collects data with a sampling rate of 30Hz and sends it via Bluetooth to the smartphone. Via Internet connection on the phone the data was sent to the server where the raw data was stored and an average breath rate for each half-minute was calculated.

In order to count calm minutes, a threshold breath rate needs to be set. The first day the default was set to 15 breaths per minute (b/min). After collecting one day of data we calculated the average breath rate during the time the person was not moving and adjusted the threshold for each individual by setting it 2 b/min lower. This assured that everybody could gain minutes by breathing calmly but not by coincidence.

Each day an online questionnaire was sent out to collect self-assessment ratings of their stress level, productivity and other psychological effects of the experiment. The participants could rate each effect on a scale of 0 (no effect) to 3 (strong effect). Not every participant filled in all questionnaires. On average 12 people filled in each one.

4.3.7 Results

On the last day of the study only 14 people could still collect the breathing data with the sensor as some participants forgot to charge or wear it.

In order to analyze the average breath rate, we separated active and passive phases depending on whether people were walking or sitting (Figure 4.2). We then compared only the data for the phases without movement. The average breath rate for all participants did not change over the week. We could not show that the calm breathing minutes per day increased. Therefore, we have to reject the first hypothesis.

To measure the psychological effect we analyzed the daily questionnaires. First we calculated the percentage of people that felt any effect at all in each category and for those people the average rating between 1 (some effect) and 3 (strong effect) for this question. All participants reported to be more aware of their breathing than before (average rating 2.12) and 82% were also more aware of their stress level (1.40). 94% agreed to take more calm breaths during the day (2.00). Further questions contributing to understanding the stress level revealed that 76% felt more relaxed when wearing the sensor (1.46), 64% could organize their tasks better (1.29) and 76% felt less tension (1.23). Consequently we can accept the second hypothesis. Due to the experiment (i) all users were much more aware of their breathing pattern and (ii) most users felt at least a bit less tense and more relaxed.

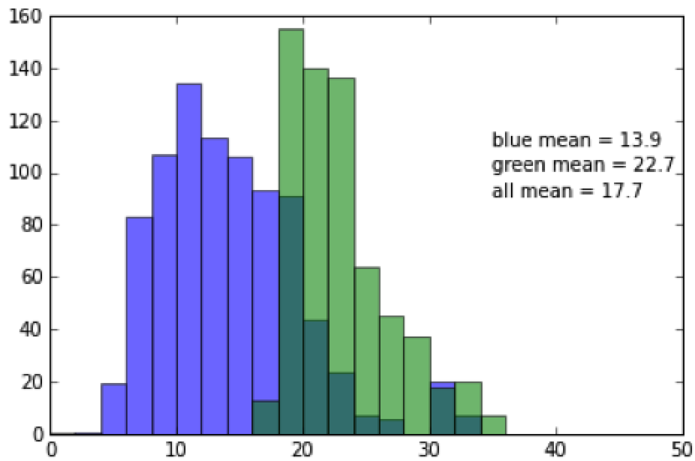


Figure 4.2: Number of minutes per average breath rate separated for breathing during rest (blue) and physical activity (green) for one subject.

4.3.8 Discussion

Especially those people already practicing yoga or doing breathing exercises prior to the experiment and who therefore better understood the importance of conscious breathing, reported to benefit from the positive effects of wearing the sensor. Irregular wearing times and the short duration of the user study made it difficult to analyze the data.

One possible explanation is that later in the week fewer people wore the sensors for less time making it difficult to compare the data.

In order to understand the possible effects better, more participants and collecting data for a longer time would be beneficial. As we collected the first data on the first day participants wore the sensor, possible effects might have occurred already during that day and no further changes occurred during the week.

However, the experiment suggests that exerting active influence on breathing patterns can cause the (subjective) stress level to decrease. Hence, a breath-monitoring device is useful to increase awareness about the state of the body and the stress level. A lower breath rate correlates with a calmer state of mind. The sensor can help to create a habit of calm breathing and integrating it in the workday [121]. Some participants wished to have more instant feedback on their breathing behavior. We took these results into account when working on the next version of the sensor and the application in order to increase the effect and user experience. In the following we describe how we proceeded in developing both.

4.4 Development of Spire

During the internship at Spire the author also worked on designing the mobile application, creating an intuitive user interface and collecting information about user interaction. These were valuable lessons also applicable when working in the startup feelSpace as discussed later in chapter 5.

4.4.1 Design heuristics

In order to design a valuable user interface for an application or a wearable device, it is important to ensure the product's effectiveness and ease of use. Therefore, these factors need to be evaluated, ideally already during the iterations of the design process. One of the most common user interface design guidelines are Nielsen's heuristics [128], briefly described as follows:

NH1: Visibility of system status,

NH2: Match between system and the real world,

NH3: User control and freedom,

NH4: Consistency and standards,

NH5: Error prevention,

NH6: Recognition rather than recall,

NH7: Flexibility and efficiency of use,

NH8: Aesthetic and minimalist design,

NH9: Help users recognize, diagnose, and recover from errors,

NH10: Help and documentation.

These design guidelines are used as an economical and quick method for evaluating interfaces while they are still under development. In this way, possible later problems can be avoided right from the beginning. The guidelines draw the attention to aspects of interface design that make their use faster, easier to remember and adaptable to specific situations.

While most of these heuristics help to improve effectiveness, another important factor that is often overlooked is to avoid users being stressed or mentally overloaded by a user interface.

TTo avoid stress for the user, design principles need to be considered when developing user interfaces for applications. Such principles can help to provide the right (amount of) information through a suitable sensory channel to the user in an intuitive way.

Consequently, stress can be prevented, boosting comprehension and supporting a positive user experience. This is achieved by reducing factors causing stress, which is crucial for a product that aims at stress reduction as in the case of Spire. Moraveji and Soesanto [123] proposed 10 design heuristics that help to avoid inducing psychological stress (M1-M10):

M1: Reveal ability to control interruptions,

M2: Reduce feelings of being overwhelmed,

M3: Acknowledge human interpretations of time passing,

M4: Use appropriate tone and emotion,

M5: Provide positive feedback to user input and events,

M6: Encourage prosocial interaction,

M7: Relieve time pressure,

M8: Choose naturally calming elements,

M9: Acknowledge reasonable user actions,

M10: Demystify the interface.

These guidelines help to consider many aspects of interfaces that might not be essential to make their use effective but that can decrease stress, e.g. by providing adequate feedback of the current status during waiting time. Reduced stress increases the likelihood of people using the interface and thus the number of customers.

4.4.2 App development

The mobile App was developed for the iPhone 5, the first model that had low energy Bluetooth required to work with the sensor. An Android app is planned but not yet available.

After various iterations of mockups and user interviews with participants of the study as well as people interested in the sensor, it became clearer what information people would like to see displayed and how to efficiently implement

(a) (b)

Figure 4.3: App design showing the amount of minutes already breathed calmly during the day. (a) Mockup versions. (b) Final version on the market.

this information (see Figure 4.3). The final application now available combines many of those features and was designed according to Nielson's [128] and Moraveji's [123] guidelines.

On the home screen a graphic display of the current breathing pattern and recent streaks (see figure 4.4a) are visible, while on the second screen a summary of the ongoing day's achievement is displayed, i.e. how many minutes the user was calm, focused, tense or active (NH8, M2) (4.4b). The first three features are calculated from the breathing patterns, while the status of "active" is calculated from the steps that are also tracked by the sensor. Tapping on one of the features opens a second screen with more details about the current streak (4.4c). Push notifications inform users about achieved streaks if desired (M1) and acknowledges the achievements or gives gentle hints on how to improve (M4, M5) (4.4d). Additionally, voice guided breathing exercises are provided to help finding more focus or calmness. A nature picture forms the background (M8) and matching colors are used for the four different types of events (NH6).

a b c d

Figure 4.4: Screenshots of the final app. (a) Main screen showing current breathing pattern and recent streaks. (b) Summary of recent achievements. (c) Details of recent streak. (d) More information about past activities and interesting facts.

When tapping on the settings symbol (NH4), one can set personal information and find help for possible problems (NH10).

4.4.3 Final Sensor

In the experiment we used the prototype that had the necessary functionality but still lacked an attractive design, essential to people liking a wearable device and using it (see Figure 4.1). Together with a designer we evaluated the needs of future customers. One important factor troubling many users of wearable devices was the need to charge it by plugging in a cable. Consequently, we made the decision to make the sensor rechargeable by induction, still a new technology at that time (NH8). The final design uses a Qi-compatible wireless charger [113]. While it is charging, a small LED light is pulsating until it is fully charged (NH1). The sensor, now called Spire stone, is designed to look like an elegant stone but is actually soft (NH8, M8). On one side it has a metallic clip to attach it to the waistband so that the soft side faces the user's body. In this position it can detect the movements of the abdomen which is used to infer breathing patterns while only the clip is visible (see Figure 4.5). The sensor vibrates to give tactile feedback when the battery is low or the positioning needs to be readjusted (NH5).

Finally, Spire could produce and ship the sensors and has been selling them since successfully.

Figure 4.5: Final version of the sensor on the wireless charging pad.

4.5 Conclusion

Breathing patterns are connected to one's mental state, especially to stress. Being aware of personal breathing habits and learning to control them, e.g. by actively reducing the breath rate, can help to decrease stress and stay more relaxed.

We performed a study with a new breathing sensor to test how feedback about breathing rates can reduce breathing rate and stress. While we could not find a decrease of the average breath rate, we could show that all participants became more aware of their breathing and most of them declared to have taken more calm breaths and to having felt more relaxed during the study period.

Together with the startup Spire we developed a new version of the breathing sensor and the corresponding smartphone app. We followed design guidelines intended to decrease stress in order to create a tracking sensor that actually helps to reduce stress by improving breathing behavior.

In the next chapter we explore another method on how to decrease stress during daily life, especially for more challenged individuals. A specific task that can be difficult is navigation, i.e. reaching an unknown location. While for some people navigation with current technological aids is feasible even though sometimes troublesome, for others it causes severe problems, and it is particularly more difficult for visually impaired people. We explore existing solutions that can support users, such as visually impaired people, especially in navigational tasks. We focus on how tactile devices can be employed in this context. To this end, we investigate how a tactile belt can use vibrotactile signals best in order to provide easily understandable guidance information.

Chapter 5

Tactile guidance in navigation

”Only he who knows his destination finds the way.”

Lao Tse (6th century BC), Chinese philosopher

For perceiving information, technology has traditionally employed vision (e.g. via screens) and hearing (e.g. via speakers, earphones). A lot of sensory input through one channel can be problematic. Especially in a context where a lot of information needs to be processed at the same time, such as during navigation tasks, where routing information, traffic signals, other travelers and much more has to be integrated. Employing the sense of touch for navigation and orientation can reduce cognitive overload by freeing other sensory capacities. In addition, using tactile feedback could be particularly beneficial for people lacking one of the primary senses, i.e. the visually impaired. A promising device is a tactile belt with integrated vibration units that can be used in navigation contexts to show directions in an easily understandable way. In this chapter, we investigate how to indicate directions best. Further, we examine whether using this device adds to already existing cognitive load. We review the potential of tactile devices and the limitations of current navigation support solutions, especially for challenged individuals. We argue that tactile belts have the potential to overcome these limitations.

5.1 Introduction

Over the past decades, technology has developed fast. Interaction with computers or smartphones through manifold interfaces have become more common and an increasingly important part of human life. This leads to a constant need to improve human-computer interaction to avoid users being overwhelmed by the complexity of the devices and applications they use. Commonly, interfaces draw on visual and auditory attention to transmit information, which can cause cognitive overload if various tasks require attention at the same time (see section 5.2.1).

Navigation, i.e. finding the correct way to reach a pre-determined location, is a common task that requires particularly much attention. Especially in unfamiliar urban environments, people need to attend to the current traffic situation, but also to street signs, names or maps to find their destination. To master this challenge, people commonly rely on technological support, such as smartphones with navigation apps or other GPS devices. Most of these devices use screens or audio instructions as means to transmit information to guide users. However, relying mostly on visual or audio output can cause problems mainly for the following reasons:

- *Internal factors* that prevent information perception because of personal limitations of the users, for example: (i) Sensory deficits like being visually impaired or deaf, (ii) overload of attention because a large variety of signals has to be processed at the same time, e.g. at a large intersection in an unknown town, (iii) divided attention because additional tasks require a lot of attention, such as monitoring the traffic or attending to others, e.g., children or (iv) they are simply not fluent in the language the information is presented in.

- *External factors* that prevent information perception for various reasons: (i) Occupation of the hands by a different task like during biking, that makes operating a display difficult. (ii) Environmental factors such as too bright or dark lighting that prevent users from perceiving the information properly on the screen. Also weather conditions (e.g., coldness or rain) complicate operating the device.

This poses a challenge especially for visually impaired people. As they lack vision, they rely heavily on their ears to be informed about their environment, making it impracticable or even dangerous to use an auditory information output when using devices for navigation support.

Tactile feedback is a valuable opportunity to avoid using visual or auditory signals when transmitting information to users. Tactile stimuli, hence, perceivable stimulations on the skin, provide an easily detectable and interpretable signal that can overcome the limitations of attention present in other channels as proposed by the Multiple Resource Theory (see section 5.2.1). Tactile feedback is increasingly used with different devices, e.g. a tactile navigation belt.

Tactile navigation belts are a prominent example of devices that provide easily understandable feedback on navigation information. A tactile belt has vibrating units all around the waist. It delivers 360°directional information via vibration at specific locations at the torso to indicate a direction. Tactile belts have been successfully applied to indicate static directions [35, 150] or guide people of different user groups along waypoints, such as pedestrians [74, 151], visually impaired people [83, 59] or the elderly [66]. Promising results showed successful navigation guided by tactile stimulation; however, many studies miss to report adequate accuracy measurements that allow decisions on the optimal use and design for such a device. Participants repeatedly reported that using a tactile belt for navigation is intuitive and does not require much attention [84, 56]. As of now,it has not been thoroughly tested how using a tactile belt affects cognitive load when applied simultaneously with other tasks. Therefore, two main questions arise: 1. What is the best way to display directions with a tactile belt to make people turn most accurately to a new direction? 2. Has using the tactile belt indeed no negative effects on cognitive load? Do tactile cues interfere with other tasks performed in parallel? To address these questions in this chapter, we perform two experiments:

1. We use a prototype tactile belt with two versions of tactile feedback to understand which design leads to more accurate responses to direction indications.
2. We investigate the interference between two tasks, using the belt and a parallel cognitive task, to verify that in fact the device is suitable to reduce cognitive load.

This chapter is structured as follows: In the literature review, we first explain the concept and benefits of sensory substitution, focusing on visually impaired people. We review current solutions they use for navigation support and their limitations that tactile devices can overcome. Summarizing prominent use cases of tactile devices, the focus lies particularly on tactile belts and the spectrum of their applicability, especially in the context of navigation tasks. We point out limitations of past research and the respective gaps we aim at overcoming with two experiments we realize thereafter. Finally, we recap the results and their implications.

5.2 Background

The common description of ”the sense of touch” often summarizes two separate senses: the kinesthetic sense and the tactile or cutaneous sense. The first provides continuous information about body movements and relative positions of different body parts to each other via receptors in muscles, joints and skin. The tactile sense provides information about stimulation on specific locations on the skin independently of the current body position. When touching or manipulating objects we use both of these senses that in combination are called haptic or tactual perception [107, 82]. The tactile sense is important for various applications using tactile feedback to provide information that is commonly delivered through other sensory channels.

Employing the tactile sense might decrease the required mental load. One main reason for mental load in the first place is the necessity to perform various tasks in parallel or attend to many simultaneous stimuli. How well various stimuli at the same time can be processed, does not only depend on the amount of input but also on the type and combination of stimuli. To calculate the effects, Wickens developed the multiple Resource Theory [181, 182] described in the next section.

5.2.1 Multiple Resource Theory

An increasing amount of time people are required to perform multiple tasks at the same time, which can cause performance decline, stress and cognitive overload. Cognitive overload is caused when (various) tasks require more cognitive capacities such as attention from a user, than is available at the current moment. Whether performance of parallel tasks declines depends on the types of tasks and the resources they demand. To predict when and to what extend engaging in multiple tasks actually leads to interference and decrease in performance, Wickens developed the Multiple Resource Theory (MRT) [181, 182].

The MRT states that information acquired from different channels can be processed in parallel without an increasing cost because different resources are available to process each. Hence, performance of different tasks at the same time depends on the combination of necessary resources for their performance. Distributing the information that needs to be processed at a given time to different input channels, can improve perception and performance. Offloading information from sensory modalities that are already highly overloaded, such as vision and audition, to less used ones such as the tactile sense can allow to process more information in parallel without a decrease in performance. The

MRT allows to predict the cognitive effort required by each of multiple tasks performed in parallel, which provides valuable hints on how to design systems and how to display information most efficiently to users.

The Prenav theory [170] proposes that the effect of external stressors on human users of technology can be reduced if the display method requires little mental resources. Hence, the theory suggests tactile displays to reduce cognitive overload.

Not always all senses are available, e.g. visually impaired people can not use devices based on visual displays. Therefore providing information through other senses can help to substitute the missing senses which is done by sensory substitution.

5.2.2 Sensory substitution and augmentation

Sensory substitution aims at providing information usually perceived with one sense through a different one. Especially difficult for people without vision is spatial orientation and the recognition of objects beyond the range of their hands or long cane [82]. One of the first corresponding experiments of sensory substitution was done by Bach-y-Rita et al. [10] in 1969 who developed the Tactile Vision Substitution System (TVSS). In this experiment a tripod mounted TV camera captured visual information. The TVSS translated the visual information into tactile patterns that was displayed via a vibrotactile array on the back of congenitally blind people. After training, they could use the information to detect and recognize simple objects in the space surrounding them. Successful implementation depended on the participants being allowed to manipulate the input. Hence, allowing them to change the directions of the camera or zoom during training led to a spatial perception of the perceived objects localized in the corresponding space outside of their reach instead of a tactile sensation on the skin [99]. In later studies, visual signals were also successfully translated to electro-tactile stimuli displayed on the tongue, the area with the highest sensibility [12].

Since then, different sensory-substitution-devices have been developed to compensate for inoperative senses such as eyesight or hearing [82, 13]. While sensory input might be limited due to physiological damage for example of the retina, the brain might still be able to 'see'. Brain structures usually processing information from a specific sense are sufficiently flexible to adapt to process different sensory input with the same contingencies, e.g. in transforming auditory information and interpret them as visual[157] or by transforming electro-tactile stimulation to visual impressions for the blind [138]. With adequate substitution devices, blind people are able to roll out actions to substituted visual input, such as grasping a drink or intersecting a ball rolling towards them [11]. This capacity can be used to develop devices that can substitute a lost sense to a certain degree and can thus arguably make blind people 'see' again [13]. When using a visual-to-audio substitution device, congenitally fully blind people can exceed the blindness acuity threshold set by the World Health Organization that defines blindness [157]. Various systems have been developed to help blind peoples orientation by substituting vision in different ways (review in [110]).

It is also possible to augment human senses by providing sensory input people normally do not perceive. While sensory substitution strives to compensate for non-functional senses, sensory augmentation provides new information human senses are not able to detect on their own [82]. One aspect of sensory augmentation is to improve a sense, like glasses or telescopes enhance vision. Another dimension is to provide a completely new sense detecting signals not normally perceived by humans. One successful tactile implementation is the feelSpace belt, a device using vibrotactile stimulation around the waist to signal information of magnetic north. Of the 13 to 30 vibration motors, the one that points north is continually activated. If the user turns their body axis, the signal adapts to the movement. After 6 weeks of training with the device, participants space perception changed substantially [124, 83, 84].

Further research on devices to help visually impaired people to navigate is described in the following subsection focusing especially on tactile solutions.

5.2.3 Navigation devices for the blind

Severely visually impaired and blind people have difficulties with navigational tasks, since navigation is highly vision based. Often, they use a white cane to avoid obstacles. Some also rely on a guiding dog or a human companion to find their way to different locations. To increase their independence and improve their quality of life, various technologies have been developed to make navigation easier and safer for them.

To solve the problem of avoiding collisions with objects, the white cane is the most common device. However, not all blind people use it. In fact, estimates range from 17% to 50% of blind people using it [111]. One of the main problems is that it usually cannot detect obstacles that are above ground level or further away than about 1 meter. To overcome this problem, various new Electronic Travel Aids (ETAs) have been proposed [63]. Some of them are sensory substitution devices (SSDs) that enhance the range by sending different signals (sonar, infrared light, laser, etc.). These devices can be pointed in various directions to measure the distance to objects out of range, which is then translated to tactile or audio feedback (for a review see [41]). One natural method implemented by some blind people is echolocation. It works by producing sounds with the mouth and then extracting information from the returning echo about present objects such as location, distance, position or even texture [92].

Golledge blind himself - made a survey with colleagues to identify preferences on necessary features of a guidance system for visually impaired individuals [65]. While they did like voice input and output, they were concerned about wearing headphones, which prevent them from listening to the environment, extremely necessary especially in high-traffic areas.

Loomis et al. [108] tested various versions of spatial display methods such as displaying sound and speech through headphones that appeared to come from the direction of the next waypoint when pointing either with the head, the whole torso or a handheld device towards the next waypoint. Participants liked the body pointing version as it kept their hands free. Furthermore, audio feedback presented through speakers at the shoulder was preferred over headphones. However, moving the whole body to point towards a direction was considered troublesome.

A different solution uses an oral tactile interface to communicate directions [161]. It displays moving patterns in four directions (left, right, forward, and backward) onto the roof of the mouth while it can be operated with the tongue from below. While front and back movements sometimes led to erroneous identification, left and right movements were almost perfectly detected. Requiring only a very low voltage (25-30V), it is an energy efficient method to provide basic directional cues to blind users while keeping their hands and ears free. However, it prevents users from speaking while wearing the device.

In unknown environments orientation is especially uncomfortable for visually impaired people as they have little possibilities to consult a map of the area they want to visit beforehand. Mental mapping helps to develop mobility and orientation skills, letting people feel safer and more likely to explore new areas. To enable blind people to explore spaces before visiting them, Virtual Environments (VE) are used. Through haptic and audio feedback they can safely get to know the place in advance [94]. Merabet *et al.*[118] could show that blind people learn the outline of a building not only when guided through an online environment simulating the same floor plan but also when they play an audio guided game based on the same floor plan even when they are not explicitly told to obtain knowledge about the layout of the building. Hence, using VE-games is a promising method to teach spatial layouts in advance to visually impaired people.

Tactile information display has the advantage, especially for visually impaired people, of not occupying the hearing sense. In the next subsection we will describe further devices that provide information via the tactile sense.

5.2.4 Tactile devices

The number of applications and devices in everyday life that use tactile feedback is steadily increasing [147, 81]. Many products available on the market use tactile stimuli, either to direct attention or to provide a more intuitive feedback on specific information. Providing information via tactile signals is gaining popularity as they are instantly noticed and often circumnavigate the bottleneck of attention when other senses are already occupied.

There are two main forms of vibrotactile displays, either 1) an entire rigid object vibrates or 2) various vibrating actuators in a spatial arrangement are used to communicate specific locations [33]. The first is mainly used in small handheld devices, such as smartphones, to communicate an event privately as they are felt just by the target person (e.g. when receiving a message on the phone) or it can be used as a warning, particularly in cars, where tactile signals are used as alarm signals to indicate distances to other cars [76].

The more flexible combination of individual vibrating units is used to indicate a specific location in space as a direction indication in navigational aids but also when a precise pattern is presented such as in the OPtical to TActile CONverter (Optacon). This device for visually impaired people converts the image from a handheld camera to distinguishable patterns on a vibrotactile pin array that can be sensed by the fingertips of the user [64].

Signals can also draw attention towards a specific location. Tapping on the shoulder via a haptic display on the back can cue attention to the correspondent side similar to real-life shoulder tapping [160]. In cars, directional vibrotactile signals have been successfully tested for various purposes such as keeping a safe distance to a leading car [76] or preventing collisions by presenting tactile cues on the side corresponding to possible danger [77]. Also, driver's mental workload could be decreased when presenting directional cues through localized tactile cues [167]. During flying maneuvers, pilots are able to decrease drifting from the main course and hence minimize navigational errors when being given simple feedback on a tactile torso display by indicating the desired direction [169].

Recently, research on supporting orientation and navigation increasingly investigated tactile feedback provided by belt-like devices with vibrators distributed equally around the waist to provide directional information in real and virtual environments. In the following section we will explore the potential of tactile belts used as navigation device for visually impaired people, but also for other user groups.

5.3 Tactile belts

Information is often transferred to humans via audio or visual presentation, but recently more vibrotactile feedback is used as described in the previous section. Especially in navigation contexts tactile feedback is beneficial as a primary information modality to avoid using other senses since vision and audition are often already occupied by other tasks. Tactile feedback provided by a belt-like device has the advantage that tactile feedback presented at specific locations around the body can indicate a direction. To present tactile stimuli at precise locations around the waist different vibration technologies are employed. For simplification, we call them vibrating units (VUs) independently of the technology used to cause the vibration and the names used in the correspondent studies. Compared to the limbs, the torso is relatively stable and therefore a good location to present directional information [34]. The concept is shown in Figure 5.1: VUs are equally distributed around the waist. One VU is activated and indicates a direction to the user.

Following the vibrating signal is intuitive and requires almost no training,

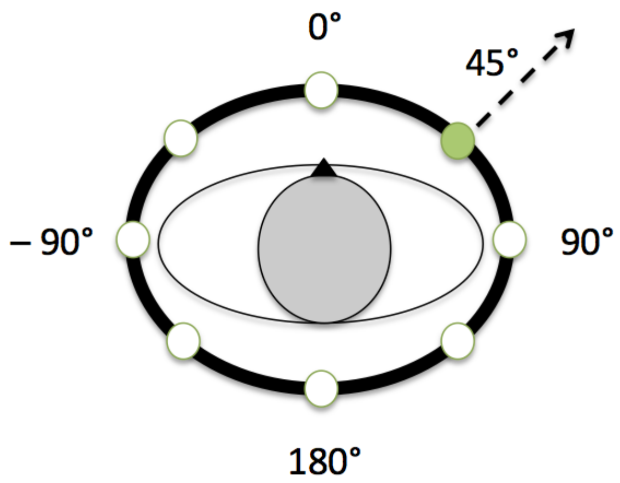


Figure 5.1: 8 vibration units (VUs) (in white) are equally distributed around the waist. The activated VU is depicted in green and indicates a direction of 45° to the user.

because the location of the directional signal is directly associated with a direction in the environment [84]. Hands, eyes and even ears stay free which is especially relevant and useful when needing the hands (in a car, on the bike, on the motorcycle) and the eyes (to monitor traffic, to enjoy nature). In fact, people start experiencing an intuitive understanding of their location and heading direction [124].

In the following, we review different tactile belts that have been tested by various research groups. Several factors are denotative for studies with tactile belts and are relevant for the type of results possible to achieve: A tactile belt is mainly characterized through the number of VUs it contains. The experiments varied in their tasks and the data they collected. The response method was in most cases either static, i.e. subjects communicated the position of the tactile stimuli but did not move, or required the subject to turn towards the indicated direction. The applied methods vary strongly, making it difficult to compare different outcomes. Often, the number of subjects is limited. In Table 5.1 we summarized a selection of prominent belt studies with relevant details.

We report the number of VUs, the number of participating subjects, the type of subjects ('normal' is used, when no specific selection criteria are given) the response type (either static or turning) and the main task in the study. Finally, we selected one main outcome specific to this study. More details to most studies are provided later in the corresponding sections.

Research in tactile guidance via tactile belts received increasing interest in recent years. Many studies investigated tactile belts in the context of navigation tasks and showed to various degrees their successful implementation. In the following, we give an overview of existing studies and their findings. Thereupon, we look at implementations and results for user groups with special needs who can particularly benefit from using tactile belts, before giving a brief overview of other application possibilities. We will then extend the researched topics by experiments building up on what we identified to be missing in previous research.

The existing research points out some limitations we will address. While successful navigation experiments have been done, no common benchmarks to measure results have been applied. Only few studies measured precisely how accurately people can follow an indicated direction or compared different cueing implementations with the same experimental settings. We describe those studies that do address turning accuracy and point out their limitations. These limitation led us to research two distinct tactile display modes to test how to increase accuracy in a first experiment.

As visible in table 5.1, most studies conduct their experiments with less than 15 participants, partly due to time consuming experimental settings [124, 84]. Mostly, belts with only 8 VUs are used, making the resolution coarse. In our experiments, we test 26 subjects with a belt with 16 VUs.

Our second experiment aims at measuring effects of mental load caused by performing parallel tasks. While MRT [181] and the prenav model [170] (see section 5.2.1) indicate that mental load could be reduced when the informa-

Table 5.1: Overview of previous belt studies summarizing the publication (name of belt or first author; year; citation), number of vibration units (VUs), number of subjects using the belt (N), type of subject (S type), response mode (either static or turning) and the task to perform. Task forms are **follow**ing waypoints, **discrimination** of indicated direction or navigation **training**. Finally, we selected one prominent result.

Belt year [Paper]	**VUs**	**N**	**S type**	**response**	**task**	**results**
ActiveBelt 2004 [165]	8	6	normal	turning	follow	correct responses
Cholewiak 2004 [35]	6, 8, 12	18	normal	static	discrimination	increasing accuracy for less Vus
feelSpace 2005 [124]	13	4	normal	turning	training	new orientation experience after 6 weeks
Van Erp 2005 [168]	8	12	normal	turning	follow	no improvement for distance indication
Pielot 2008 [136]	6	16	normal	static	discrimination	accuracy increased by interpolating vibration intensiyity between 2 Vus
Jones 2008 [81]	8	10	normal	static	discrimination	98% correct recognition of active VU
Tactile Wayfinder 2008 [74]	6	7	normal	turning	follow	subjects stay within 15m of path
Tactile Wayfinder 2010 [135]	12	14	normal	turning	follow	requires less attention than common GPS
Grierson 2009 [66]	4	9	elderly	turning	follow maze	performance higher than for speech based system
Grierson 2011 [67]	4	11	dementia	turning	follow indoor	successful guidance for dementia patients
TacNav 1 2011 [150]	8	16	normal	static	discrimination	better than tactile back array
TacNav 2 2013 [151]	8	12	normal	turning	follow	faster than other GPS device
feelSpace 2012 [83]	30	1	blind	turning	training	improved sense of security
feelSpace 2014 [84]	30	7	normal	turning	training	new orientation experience after 7 weeks
Cosgun 2014 [38]	8	15	normal	static	discrimination	Continuos vibration lead to best results
Faugloire 2014 [56]	8	12	normal	turning	turning	Tactile indication leads to better accuracy than speech
Li 2015 [103]	4	18	normal	static	discrimination	Comparison of frequency, amplitude, duration
Flores 2015 [59]	8	10	blind	turning	follow	Path following more precise than with audio guidance

tion necessary to perform parallel tasks is perceived through different sensory channels, only few studies precisely measured the correspondent effects in the context of tactile belts. We give a short overview over the existing studies and point out the gap we address.

5.3.1 Tactile belts for navigation

One of the first tactile belt prototypes was build by the feelSpace group [124] and it provided cardinal information via localized cues. 13 VUs were placed equally around the waist. The VU closest towards north was vibrating continuously, providing a constant feedback of cardinal direction (4 belt and 4 control subjects). Within 6 weeks of training with the belt, in particular performance in an outside navigation tasks improved. In half of the participants, a qualitatively new sensory experience emerged so that they intuitively ”knew” were North was and how they were orientated in space. A later experiment of the same group used a new prototype with 30 VUs [84]. 9 participants wore the belt during the duration of 7 weeks and used it while training navigation tasks. The belt facilitated navigation and stimulated using new navigation strategies.

Another early belt prototype was the ActiveBelt with 8 VUs equally distributed around the waist [165]. 6 participants were able to identify cardinal direction correctly and follow the direction indicated by the belt.

Van Erp et al. [168] tested a further belt with 8 VUs. 12 participants followed a path with 6 waypoints . Different methods of encoding the distance to the upcoming waypoint in addition to the direction did not change performance, suggesting that distance information may not be relevant for efficient guidance. This belt was later also tested in the context of challenging military tasks [51]. When testing the tactile display in comparison to standard GPS devices and combinations of both for navigation tasks in difficult terrain and with an additional task requiring attention, they concluded that the tactile navigation display has large application potential especially when subjects are under conditions of heavy visual and cognitive load.

The Tactile Wayfinder is a belt with only 6 equally distributed VUs. To indicate an intermediate direction, two neighboring VUs were activated at the same time [74]. On an open field, 7 participants followed the signal along waypoints.

In 99% of the time they stayed within 15m from the target. While authors identified these results as success, leaving the core path up to 7.5m to either side is potentially dangerous in a urban environment. A later version included 12 VUs [135] and was used to convey information not only of the next but of the next two waypoints. To achieve this, a sequence of two outputs was presented every 4s that represented the next two waypoints respectively. In a study with 14 subjects comparing the Tactile Wayfinder with a standard GPS device, the tactile belt caused more navigation errors but required less attention as subjects remembered the environment better. However, no alternative waypoint presentation patterns were tested.

TactNav is a belt with 8 VUs that was first compared to a tactile array displaying information on the back [150]. TactNav outperformed the back array in a direction discrimination task with 16 subjects. Further on, the performance of TactNav was compared with a navigation app on a smartphone [151]. In a field experiment with pedestrians navigation accuracy was similar in both cases, but route completion time was significantly faster with the tactile belt.

Other studies investigated how to use a tactile belt to provide directional information to drivers in a car. Boll et al. [18] developed vibration patterns that indicate the intended turns in advance. Asif et al. [8] developed rhythmic patterns to additionally show distance information to car drivers via a tactile belt.

5.3.2 Navigation belts for the blind

To test a minimalistic approach of tactile feedback, Marston et al. [114] compared an audio and a tactile binary cue setting. 8 blind subjects wore a head mounted compass and received either auditory or tactile feedback presented via a single VU placed on the arm to indicate whether they are within 10°of the desired path. In both conditions, subjects were able to successfully follow the path. They proposed a multivibrator display to overcome the limitations occurring when switching from walking towards one waypoint to the next one, where participants in this study had to actively turn until they find the next direction. However, they did not test the alternative display.

In a case study, one late blind subject trained with a tactile belt with 30 VUs, resulting in improved performance in various navigational tasks. Most of all, the tactile device enhanced the feeling of security especially when exploring unfamiliar environments, which points out the opportunities such devices present for self-guided and self-depended navigation of visually impaired people [83]. In this case, information about cardinal north was displayed continuously through a tactile stimulation on the correspondent location on the waist.

Flores et al. [59] used a belt with 8 VUs or audio feedback (speech) to guide 10 congenitally blind subjects along 6 paths. The belt indicated directions by activating the correspondent VU pointing towards the next waypoint. Thus, it created a moving signal to indicate the need for turning (e.g. a signal moving from the front towards the right to indicate the need for a right turn). With the tactile feedback, participants were slightly slower but followed the path more precisely. In general they preferred the tactile version to avoid impairing their hearing sense.

5.3.3 Navigation belts for the elderly

A wearable tactile belt with 4 VUs was used to guide elderly people through a maze. When compared to younger adults, they made significantly more navigation errors when guided by speech but followed the tactile signal without problems even when the route complexity was increased [66]. The same device was also successfully used by persons with mild dementia [67].

5.3.4 Other use cases

One important advantage of a tactile belt is the possibility to present stimuli at specific locations. This is not only useful when indicating directions for navigation purposes. The belt can also be combined with an obstacle detecting system. Johnson and Higgins [79] developed a prototype with 14VUs. Cameras attached to the devide detected the objects around the user, the location was transmitted to the user via vibrotactile feedback at the correspondent locations. Cassinelli et al. [31] used a band around the head with 5 sensor-vibrator modules that could detect obstacles at the height of the head and provide correspondent tactile stimulation. Blind users felt safer and more independent with the band and their anxiety related to navigation decreased.

Obstacle detection is also an important task for robot locomotion. A tactile belt was successfully used in the context of remote controlling a robot. To represent the distances to objects from the robot, localized tactile feedback was presented to the person controlling the robot's movements through the interface of a tactile belt [164].

In general, tactile cues are a good way to attract attention towards a specific location, which can also be used in other scenarios. Borg et al. [19] successfully indicated the source of sounds to deaf people via a tactile belt. Buchmann et al. [28] in contrast used tactile cues in the context of a Virtual Environment to indicate the direction of a target the user should turn to. In comparison to visual and auditory hints users performed best with the tactile belt. Ferscha et al. [57] used a tactile belt to successfully show workers the location of dangerous obstacles close to their path.

5.4 Tactile direction display

So far, we summarized research on tactile devices and their potential for transmitting information particularly when other senses are unavailable or used otherwise. We reviewed experiments with tactile belts that are specifically useful for navigation applications, not only but particularly for visually impaired people, because tactile belts can overcome the limitations present when other senses are compromised. to receive guiding instructions.

In various studies, the effectiveness of tactile belts for providing guiding instructions in form of directional tactile cues, has been shown. However, it remains unclear how to indicate directions best to users to achieve the highest accuracy for tactile guidance. In the following, we review previous methods of displaying directions and measuring the consequent responses. To close the gap we find in prior research we design two correspondent experiments.

In the first experiment, we compare turning accuracy for different tactile display modes. Turning accuracy denotes the precision with which the tactually presented direction is estimated by the subject [56]. We test how accurately subjects turn towards the direction that is displayed via the tactile belt. Therefore, we call the result turning accuracy achieved in a turning task.

With the second experiment we examine the effect of mental load on parallel tactile guidance to test whether performing both tasks in parallel compromises performance.

5.4.1 Location of tactile stimulation

In previous tactile belt research, the position of the tactile stimulation on the body changes in order to indicate different directions. A target location is displayed by activating the VU that lies in the direction of the target. VUs are normally equally distributed around the waist and each VU has an equal range of directions that it represents [165, 124, 168, 74]. Hence, the range of directions of each VU depends on the number of VUs integrated in the belt. Consequently, the fewer VUs a tactile belt has, the lower the resolution of angles it can indicate, e.g. when having 6 VUs, each represents directions in a range of 360°divided by 6 equals 60°.

Pielot et al. [136] tested a different approach to display directions with a belt that was limited to 6 VUs. When displaying a direction between two VUs, the two neighboring VUs are activated and the exact angle of the direction is encoded through different intensity levels of their vibration. The one closest to the target direction vibrates with a higher intensity. For 15 subjects, this approach increased the accuracy of direction detection significantly, but at the cost of slower reaction time. Participants found it more difficult to interpret the signal compared to discrete directions displayed with only one activated VU.

To evaluate the best way to provide directional information to blind people Cosgun et al. [38] considered not only directional cues but additionally moving vibration patterns to directional signals. However, continuous vibration of a single VU to indicate a direction leads to least recognition errors, while two intermittent pulses were slightly preferred.

5.4.2 Accuracy of perceiving directions

In research, an important factor to judge effectiveness of tactile belt stimulation is the method to measure the accuracy of response. Two types of studies can be classified according to the required response requested from the participants. The first type requires a stationary response and focuses on the localization of the signal. Participants either use a handheld rotary dial to manipulate a cursor [166], press buttons arranged corresponding to the stimulus site [34] or draw the indicated direction on paper [103]. The second response type requires that participants turn actively towards the direction they want to indicate [56] or additionally walk towards it [135, 168, 59]. While measuring stationary responses, contrary biases have been found shifting the perceived signal either towards the sides of the abdomen (i.e. away from the navel or spine) [34] or towards the middle of the body (i.e. towards the navel in the front and towards the spine in the back) [166]. It seems that indicating a direction via body movement is an ecologically more valid method. Haber et al. [71] showed that the method of indicating a direction has an influence on the accuracy of the perceived location. Using the whole body or parts of it to indicate the direction, leads to the best results. Many studies proved the effectiveness of tactile direction displays. However, often users performance was either not reported completely or the study did not include accuracy measurements of the body's orientation movements. Other responses were evaluated such as discriminating which specific VU was activated. [35, 150].

Most of the devices mentioned in literature used 8 equally distributed VUs around the waist and mapped equal angles to each [165, 168, 150, 38, 59]. In other words, the tactile belts most often have a resolution of 45°per VU. To test the accuracy of turning performance, most studies only tested whether the "correct angle" is perceived. Cosgun et al. [38] tested the turning accuracy for different stimulation versions. Participants walked around and verbally named the VU number they felt vibrating. When the vibration signal was presented continuously on the same VU, the turning accuracy was highest (9° deviation).

Heuten et al. [74] tested a belt with only 6 VUs. In this case each VU represented an angular range of 60°. Therefore the expected deviation between actual and presented directions lies between 0°- 30°with an expected average deviation of 15°. In fact, in experiments with 13 participants, the direction indicated by the signal produced by the belt was perceived with the expected median deviation of 15°. Pielot et al. [136] used a similar belt but the real average error for 16 participants was higher with an average deviation of 19.4°. The authors proposed as possible explanation the difficulty of mapping the perceived signal to the response format, a visual circle on a screen.

However, in real navigation tasks, 45° angles are quite inexact, especially if the user has to navigate along a route that is not visible (e.g. because of darkness or visual impairment) or where the destination is not clearly recognizable (e.g. in an open field, a forest or on a big parking lot). Thus, a better resolution is desirable.

To reduce the recognition error, the number of VUs could be increased. But this number has its limits not only for economic reasons, but also because different tactile stimulations can only be differentiated on the skin when the distance between two stimulation sites has a minimal distance of at least about 35mm (depending on the location around the waist) [180].

A further factor having an influence on perception of tactile stimulation is the vibration rhythm of the tactile signal. In previous tactile guidance studies, the vibration rhythm was chosen mostly arbitrarily, either with constant rates [150, 66, 59] or variations in order to code distance information [168, 18]. In the most cited tactile guidance paper, Van Erp et al. [168] used pulses in 1s intervals and recommended a feedback frequency of at least one pulse every 4s. However, the effectiveness of this recommendation has received little validation [56].

The chosen vibration rhythm determines how this information can be used to control turning movements. When the vibration rhythm is high (several bursts per second or even a continuous signal) the location of tactile stimulation can be updated in accordance to the rotation of the user's body. Such a relation between movement and tactile feedback creates an action-perception coupling similar to natural sensory perception (e.g. updated perceived image during body rotation) [130] that has been found in users of tactile devices in previous studies [124, 83].

Faugloire and Lejeune [56] directly compared whether this more natural approach for presenting tactile information enhances turning accuracy in contrast to a short initial direction indication without updating the signal during movement. In an experiment with an 8 VU belt, 12 participants turned towards the indicated direction and turning accuracy was measured. A continuously updated signal led to higher accuracy (absolute turning error AE = 10.1°) than compared to a short initial burst (AE = 13.6°). The difference remained the same when performing the experiment in the dark. The findings show that action-perception coupling supports the effectiveness of tactile guidance. Hence, it is useful when the tactile signals are reciprocally related to the movement of the user. This condition was also favored by participants.

Our goal is to reach the highest turning accuracy possible via tactile guidance. The described experiment [56] reports with AE = 8.9°one of the lowest errors for turning accuracy. Therefore we base our study on their study. To use the benefit of action-perception coupling, we use a continuous vibration in our experiment so that the tactile signal was constantly updated according to subject's movements. We used a belt with 16 VUs to increase the expected accuracy. We then compared two different display modes, i.e. different mappings of directions to VUs, to measure their influence on turning accuracy.

5.4.3 Mental load

The Multiple Resource Theory (section 5.2.1) proposes that providing information through different sensory channels can reduce mental load, especially when tasks are performed in parallel [182]. A meta-analysis of tactile applications shows improvements for workload and performance when adding tactile to existing visual cues [50]. Therefore, tactile devices can provide simple information additionally to vision or audition, reducing mental load.

Participants in various studies repeatedly reported that using the tactile belt for navigation required less cognitive effort or attention than other methods [84]. A few studies investigated the effect of using a tactile belt on mental load. Elliott et al. [51] measured self-ratings on mental workload and found that tactile navigation displays can outperform visual displays when cognitive and visual load is high. Davis [43] measured self-rated mental load for 7 different navigation devices but did not find any specific advantage of the tactile display mode. However, in contrast to a visual personal pedestrian navigation device, Pielot et al. [135] confirmed that pedestrians using the tactile belt attended more to their surroundings. Dorneich et al. [45] compared performance of 9 subjects under low and high cognitive load conditions. They performed various cognitive tasks in parallel to a navigation task either supported by a tactile belt or visual based navigation device. Subjects performed better with the tactile belt when cognitive load was high.

Especially for visually impaired people performing navigation tasks requires high levels of attention. In a case study with one late blind subject, Krcher et al. [83] tested the effect of an additional mental task on the capacity to walk a straight line. Walking a straight line has a high relevance for blind people for example when crossing large streets and simultaneously using the ears to monitor traffic. When performing the mental task, performance was significantly better when using the tactile belt than without it. In this study, the tactile belt continuously indicated north, hence the signal was not necessarily felt in the front but at a consistent location around the waist. In a more complex navigation task, an additional mental task caused worse performance independently of the support of the belt. However, after 5 weeks of training with the tactile belt, performance was higher when using the belt, suggesting that the signal was integrated and could be used unconsciously [83]. Klatzky et all [88] tested mental load on a navigation task without vision when guided by either virtual sound or language. When under mental load, navigation performance decreased for the verbal condition but in the virtual sound condition, performance dropped only for the first trial. In the consecutive trials, subjects performed again equally well compared to the navigation-only task, suggesting that very short adaptation or training is enough to avoid interference between modalities. .

While some studies evaluated subjective mental load via personal ratings, only a few studies measured concrete performance and those predicate their results on only few subjects (9 in [45], 1 in [83]). In their conclusion of their recent paper Faugloire and Lejeune [56] explicitly pointed out the necessity to test the effect of mental load on a tactile guidance task. In our experiment, we aim at closing this gap. We test 26 subjects and measure both, the performance in the cognitive task and in the tactually guided task. Both tasks are performed by themselves as well to allow drawing conclusions on mutual influences the tasks might have on each other.

In the present literature, we identified various aspects expected to increase turning accuracy during tactile guidance. With the two experiments we conducted, we aim at closing the gaps we identified in recent literature.

As vibration rhythm we used a continuous vibration that proved of value because it allows for action-perception coupling [124, 84, 38]. With the first experiment, we aim at quantifying how well the directional display of a 16 VU tactile belt can be understood.

As we saw in the last section, still no clear results show the effects of mental load on a tactile guidance task. Therefore, in the second experiment, we investigate the effect of mental load on the turning accuracy.

5.4.4 Experiment 1: Turning accuracy

In our first experiment we quantify the turning accuracy achieved by two different display modes: precise action-perception coupling versus emphasizing the target direction.

When each VU indicates directions in the range of 22.5° (360°/16 VUs), an maximal error of 11.25° in each direction is possible with an expected average error of 5.6°. However, this would only be true for an optimal situation in which the user precisely detects every signal change and differentiates well between the closely spaced VUs. It has been shown that the accuracy of distinguishing between different VUs decreases when more VUs are present (for 12, 8 and 6 VUs) [35].

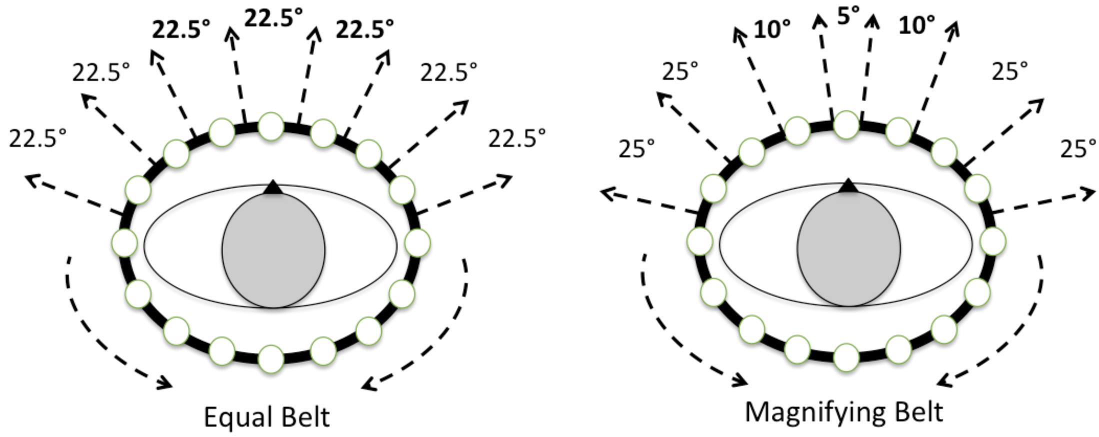
But with fewer VUs, the amount of possible directions that can be displayed declines and accordingly turning accuracy decreases due to the ample range of directions represented by each VU [136, 74]. Many studies use a tactile belt with only 8 VUs (see table 5.3). Yet, how precise directions can be detected with a tactile belt with 16 VUs has not been tested so far. During the first experiment, we test the turning accuracy with 16 equally spaced VUs around the waist and compare the effect of two different display modes on the turning accuracy.

The first mode we call the equal belt (EB), where the 16 VUs represent an angle of 22.5° each. Hence, one VU is active when the target direction lies within its range (see figure 5.2 a). This allows a precise action-perception coupling, i.e. body movements are leading directly to correspondent tactile feedback. The tactile feedback changes in a consistent manner when turning the body and with a consistent difference between VUs, independently of the direction.

However, tactile discrimination accuracy is highest close to the navel [35, 166]. Therefore the second display mode, the magnifying belt (MB), displays directions differently: The VU directly in the front close to the navel covers a range of only 5°, the two next to it, one on either side, cover 10° each while all others cover 25° (figure 5.2 b). Here, a less direct action-perception coupling is present. The signal still wanders when the user turns, but the tactile feedback changes faster during turning when the target direction lies nearly in front of the person, compared to all other directions. We hypothesize that the higher resolution in the front leads to increased turning accuracy.

5.4.5 Experiment 2: Mental load

In a second experiment, we tested the tactile belt with an additional mental load. In an application case, pedestrians during navigation have to fulfill various tasks in parallel, such as monitoring the traffic while finding the way to their destination. Hence, a device supporting navigation should be easy to use and not disturb other tasks performed simultaneously. To learn about the influences that using a tactile belt and performing mental tasks have on each other, we tested whether the turning accuracy stays the same when using the tactile belt



(a) (b)

Figure 5.2: Schema of the distribution of the represented directions among the 16 VUs for the two different display modes: (a) The Equal belt (EB) and (b) the Magnifying belt (MB).

while additionally adding cognitive load through a mental task. We expect the two tasks not to effect each other like the Multiple Resource Theory proposed (see section 5.2.1) [182]. To induce mental load, we used a version of the 7-backwards task that was used for the same purpose by Krcher et al. [83].

5.4.6 Hypotheses

In conclusion, our hypotheses for the two experiments are:

1. When using a tactile belt with 16 VUs were each represents directions in the range of 22.5° (EB), an average turning accuracy of at least 5.6° can be achieved.

2. When the range of the 3 frontal VUs is reduced (MB), the accuracy of turning performance should increase to an average of around 1.25°.

3. We expect the reaction time to be the same in both conditions. This is because in both cases, a continuous vibration signal is presented and action-perception coupling is present.

4. When performing the turning task and the mental task simultaneously, we do not expect a performance difference compared to performing each task individually.

5.5 Methodology

5.5.1 Participants

26 participants (11 female, 15 male) aged between 19 and 55 years (mean = 25.0, $\pm$ 6.8) took part in the study. 5 had already participated beforehand in a different study with a tactile belt. All had normal sensory functions. In the beginning, each participant received written information about the experimental procedure and signed an informed consent statement.

The subject was seated on a swivel chair with at least 80cm free space in each direction. The tactile feelSpace belt was put on and closed in the back. The belt is a research prototype from the feelSpace group that contained 32VUs, but only every second one was activated. Hence, it used 16 equally distributed vibro-tactile units (VUs). Each VU is a standard pancake vibro-motor like those used in smartphones vertically positioned.

The belt was connected via Bluetooth to a Smartphone (Nexus 6) that was fixed on the arm rest of the chair and pointed towards the front. In this position it stayed parallel to the participant while he[1] performed the turning task. The belt was operated by a smartphone app that saved data every 100ms. The data contained both, the current direction indicated by the belt (D1) and the actual heading direction of the participant (D2) together with a time stamp.

[1]For easier reading we write about subjects as 'he', but mean both our male and female participants.

5.5.2 Procedure

After putting on the tactile belt, the participant sat down on a swivel chair and the experimenter explained in detail how the belt works. Then the experimenter started a compass mode in which the belt continuously indicates north. The subject took the smartphone in his hand and received a visual feedback indicating north. He was asked to familiarize himself with the sensation of the tactile signal while swiveling on the chair to experience the wandering of the signal. Finally, he was asked to stop turning when heading towards north, i.e when the tactile signal was displayed exactly at the VU closest to the navel. This served both for familiarizing the subject with the task but as well for letting the experimenter control that the belt is correctly placed and the accuracy of the smartphone's compass was not compromised.

Further details were explained if the subject had questions. Then, the subject was asked to close his eyes for performing the task. The reason was that the subject should not be distracted by movements of the experimenter or other details of the room. All participants closed their eyes as asked and only opened them between trials (monitored by the experimenter).

5.5.3 Experiment 1: Turning task

The experiment consisted of 4 trials, 2 using the MB and 2 the EB in alternating order. The beginning condition was counterbalanced across subjects. Each trial started with 3 bursts on the side of the belt during which the app calibrated itself followed by a vibration in the front for 5s. During this initial phase, the participant did not move. The signal then deviated to a new predefined angle and the correspondent VU vibrated for 10s. Every time the signal changed, the participant rotated on the chair so that he was aligned with the signals direction with the frontal VU vibrating again. The angle changed 9 times resulting in a total length of 95s per trial (5s initial vibration plus 9 directions á 10s). Each trial contained 9 pseudo-randomized angles: at least 4 direction changes leftwards and 4 rightwards, no more than 2 changes towards the same direction in a row and at least 2 small changes (between 10°and 90°) and 2 big changes (between 100°and 180°) in each direction.

In total, each possible direction in steps of 10° between 10° and 180° was represented twice, once to the right and once to the left, distributed across the 4 trials (36 directions in 4 trials = 9 directions per trial).

5.5.4 Experiment 2: Mental task

After finishing the first experiment, participants continued directly with the second one.

Each subject performed a mental task in 4 trials. In 2 trials, only the mental task was performed and in 2 trials, the mental task was simultaneously performed with the rotation task (using EB) from the first experiment. During the mental task, the subject continuously subtracted 7 from a 3-digit number initially announced. 4 numbers (451, 519, 778, 982) where randomly assigned to the 4 trials so that each subject started with the same 4 numbers in a differing order. In this way, we avoided varying difficulty in the mental task between subjects. When performing the turning task together with the mental task, the experimenter told the initial number right before starting the belt simulation and the participant announced the final number he reached as soon as the simulation was over. In trials where subjects performed only the mental task, the subject sat still on the chair until the experimenter announced the end of the trial period after 95s. The subject announced the final number he arrived at. The experimenter noted the results in both conditions. The trial versions were performed in alternating order. The beginning condition was again counterbalanced between subjects.

5.5.5 Qualitative scores

After the experiments, participants completed a questionnaire asking for qualitative feedback. They were asked to score easiness (Did you find the task easy to perform?), intuitiveness (Did you find the directional information intuitive?), and perceived accuracy (Do you think your responses were accurate?) for both display versions. The rating scale ranged from 1 (very difficult, not intuitive at all, very inaccurate) to 6 (very easy, very intuitive, very accurate). Finally, the participants were asked whether they preferred one of the 2 belt conditions. Altogether, the experiment took about 30 minutes to complete.

Additionally, subjects were asked whether they believe to have performed better at the mental task either with (1) or without (5) the parallel turning task on a 5-point Likert scale.

5.6 Data analysis and results

The app saved one file for each trial, logging both D1 (the direction indicated by the belt) and D2 (the actual heading position) for every 100ms (see Figure 5.3 for an example of the data). To extract the absolute error (AE), we measured the difference between D1 and the mean D2 during the last second of the presentation of D1, when the participants had usually reached a stable position.

To evaluate the reaction time (RT), we extracted the time the participant needed from the onset of a new D1 until he turned more than 3° away from his previous stable position. With 3° we assured that we did not capture artifacts caused by inaccuracy of the compass or micro-movements of the participant. We did not evaluate the time used to arrive to the new stable position because i) this time is dependant on how big the angular change is and ii) we gave no clear instructions on how fast people should turn. Some participants moved slowly to avoid feeling dizzy from the turning movement. As each D1 was displayed for 10s, participants usually had enough time to turn. We visually checked the

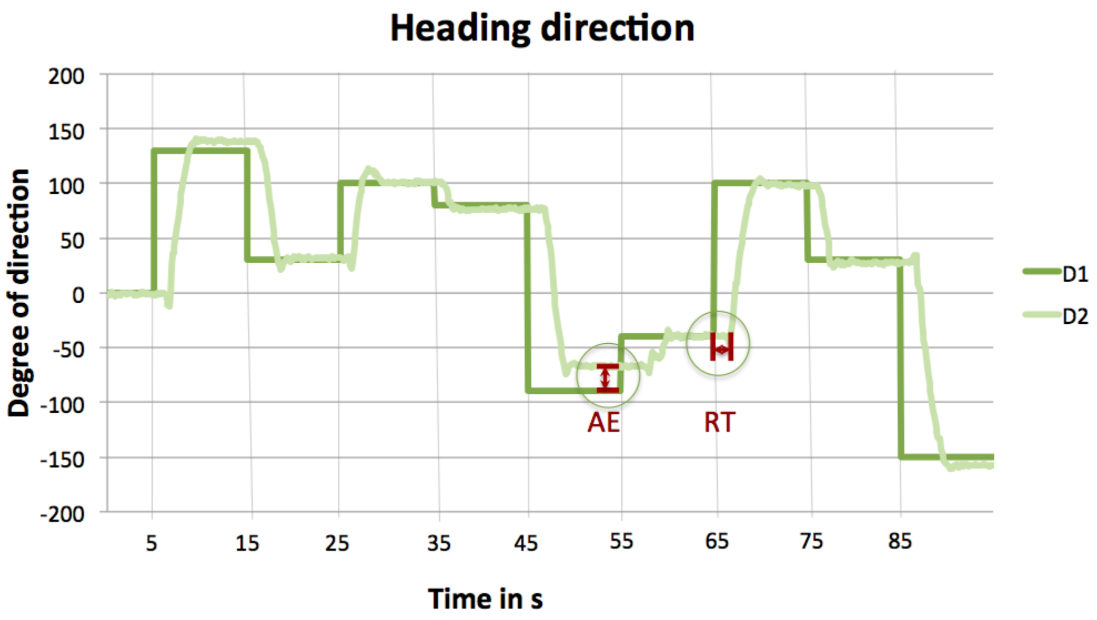


Figure 5.3: D1 shows the direction displayed on the belt while D2 shows the real movement of the participant. Changes upwards indicate turns to the left and changes downwards turns to the right. The absolute error (AE) between D1 and D2 is the mean distance between the two during the last second of presentation.

data to detect cases in which people did not arrive in time before a new D1 was displayed; in this case, we removed the data point. For some participants, a VU failed to work. If therefore he missed a direction (because no vibration was felt), we also excluded the event. If, in contrast, the subject noticed right away that the vibration was missing, he turned slightly until the signal was felt at another VU again and completed the task. In this case we included the data for measuring the AE but not for RT. A few data points had to be removed because the compass was malfunctioning and caused erroneous direction displays. Two subjects had to be excluded because no stable signals had been displayed. In total, 8.6% of the data was removed.

In case of the secondary task (subtracting 7 from a random number), for each trial, we collected the number the participant began with and arrived at. From that, we calculated the number of iterations he performed (N7) and whether the final number was calculated correctly. If the final number was not a correct result, we rounded to the closest correct answer. Participants achieved on average N7= 21.7 (±7,8).

5.6.1 Display modes

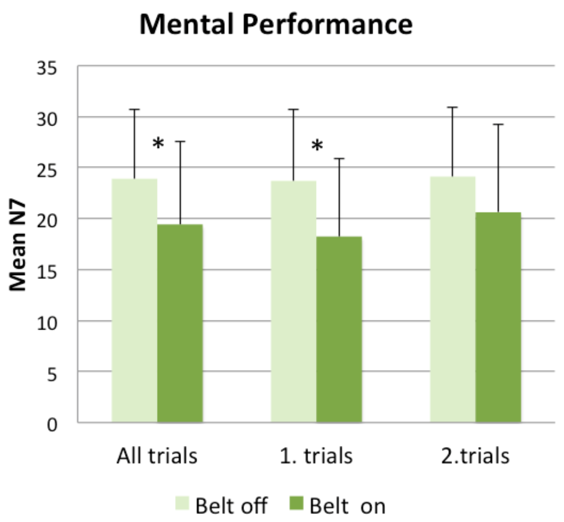
According to the first hypothesis we expected a mean error of 5.625°in the EB condition. Participants achieved a significant better result with 4.912, tested with a one-sample t-test ($p = 0.0122$). We compared the mean AE for each subject between the two belt settings EB and MB with a paired two-tailed t-test. The mean AE was almost exactly the same in both conditions with around 4.9°. Accordingly, no significant difference could be found ($t(46) = -0.0036, p = 0.997$). Also, for mean RT no difference between the two settings was detected: $t(46) = -1.1696, p = 0.2541$. The mean RT for the EB condition was 1.35s and 1.43s for MB respectively.

5.6.2 Effects of the turning task on mental performance

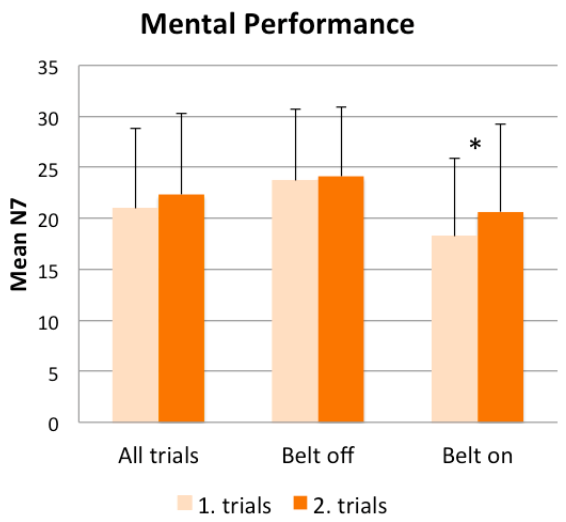
No difference of N7 was found for the 4 initial numbers confirming that the difficulty of the mental task was similar, independent of the starting number. The error rate with respect to the final number was similar for both trial versions. When performing the mental task together with the turning task, 60.3% of the final numbers were correctly calculated. Respectively, 60.4% were correctly calculated when performing the mental task alone.

To test whether there was a negative effect on the performance in the mental task when simultaneously doing the turning task, we compared N7 for both conditions, the mental task alone (”belt off”) and both tasks together (”belt on”). Every participant performed two trials for each condition. When comparing the mean N7 result with a two-tailed paired t-test, the difference was significant ($t(46) = 4.5889, p < 0.001$). In the ”belt on” -condition they reached N7 = 19.3

while they reached N7 = 23.9 when only performing the mental task. When separating in blocks of the first two and the second two trials, a difference is clearly visible. In the first block, the performance differs significantly between conditions with a difference of 5.5 ($t(46) = 5.5189, p < 0.001$), in the second block, the difference of N7 is lower (3.5) and not significant anymore ($t(46) = 1.3281, p = 0.1972$). When comparing the mean of the complete first block with the second ("All trials") no significant difference can be seen (see Figure 5.4 a).



(a)



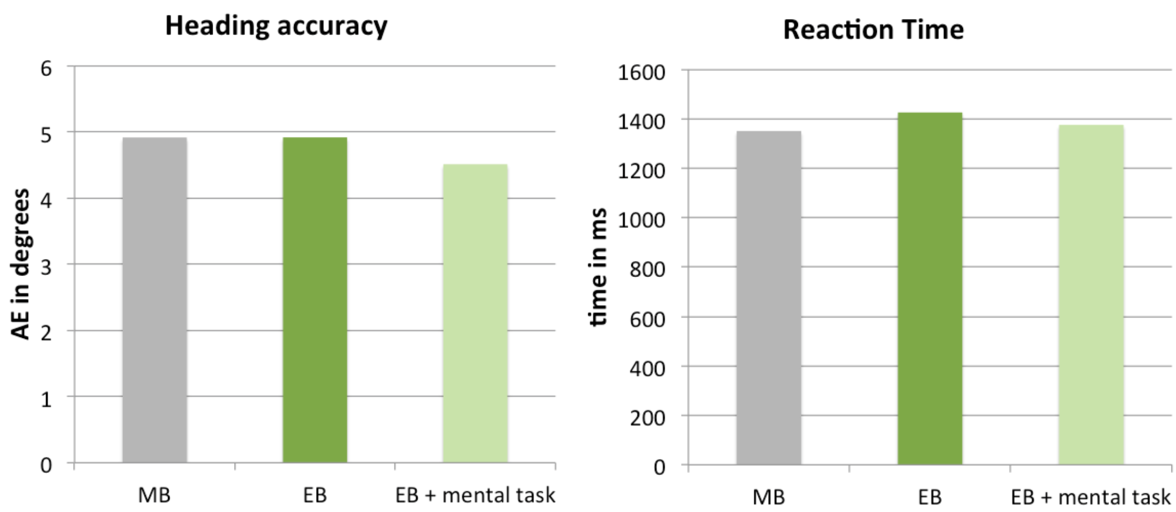
(b)

Figure 5.4: Performance in the 7-backwards task: Mean number of iterations subjects subtracted (N7) with performing a turning task with the tactile belt (Belt on) and without (Belt off). We compare average performance for all trials and separate N7 for the 1. trials and the 2. trials. First we compare N7 depending on (a) the belt on or off condition and then (b) separated for the first and second blocks of trials. * denotes a significant difference. Error bars indicate the standard deviation.

In general, no significant learning effect was found, i.e. participants did not improve over time in the task. When comparing the first and the second "belt off" trial, no difference was found ($t(46) = 0.3796, p = 0.7078$). In contrast, when comparing the two "belt on" trials a significant N7 improvement of 2.3 was detected ($t(46) = -2.4155, p = 0.0241$) (Figure 5.4 b).

5.6.3 Effects of the mental task on the turning performance

We tested whether performing the mental task in parallel to the turning task had a negative impact on the reaction time and the turning performance. Therefore, we compared RT and AE to those from the EB condition that used the same belt setting but without an additional task. The turning accuracy is similar in both conditions with no significant difference in AE ($t(46) = -1.2425, p = 0.2266$). Also the reaction time does not slow down when adding the additional task ($t(46) = -0.5385, p = 0.5954$). Both comparisons are displayed in Figure 5.5.



(a) (b)

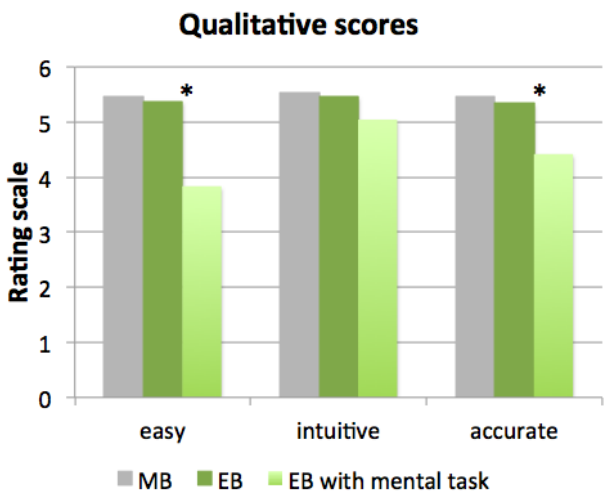
Figure 5.5: Comparison of (a) turning performance and (b) reaction time for the two display modes and when adding the mental task. The mean AE and RT show no significant difference between conditions.

5.6.4 Qualitative data

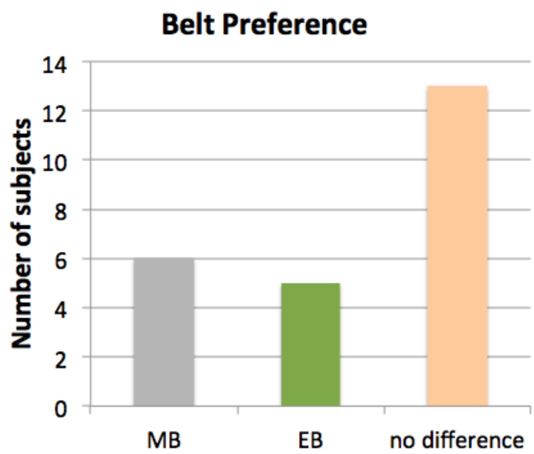
When comparing the qualitative data we found no difference between the two belts MB and EB. Both versions were rated equally well with average ratings between 5.3 and 5.6 for all three categories, hence subjects perceived the use of the tactile belt as easy, intuitive and expected their answers to be accurate (see

Figure 5.6 a). No clear preference for one of the display modes could be found (see Figure 5.6 b) While 13 subjects had no preference, 6 preferred MB and 5 EB.

When comparing the ratings for the combination of turning task and mental task, only the perceived intuitiveness of the tactile signals was still the same. Participants perceived the task as significantly harder and believed to react more inaccurately. The significant differences are denoted with a * in Figure 5.6 (a).



(a)



(b)

Figure 5.6: Qualitative data: (a) no significant difference has been found for rating of easiness, intuitiveness and accuracy for the magnifying (MB) and equal belt (EB). When comparing EB with the ratings for the mental task a significant difference is present for perceived easiness and accuracy (marked with *). (b) There is no clear preference for the display mode .

5.7 Discussion

5.7.1 Turning accuracy

We could show that the turning accuracy with a 16 VU belt is very high with an AE of only 4.9°, which is significantly lower than the estimated 5.6°. Therefore, we can accept our first hypothesis and confirm the high turning performance precision.

When comparing the two tactile display conditions, we find no difference in turning accuracy. We expected AE to decrease since the smaller range of directions represented by the frontal VU should lead to a higher turning accuracy. This was not the case. Subjectively, the two display modes were perceived to be similar and no clear preference could be found.

One possible explanation for the fact that we found no difference between the conditions could be the strong effect of action-perception coupling in both cases. When subjects turn, they quickly know how the tactile signal changes according to their body movement. When they feel a new direction indication and start to turn, they can already estimate how far they need to move. The subjects acquire a feeling for how far exactly they have to turn for the signal to change to the next VU. Consequently, they do not stop as soon as the signal switches to the frontal one. Instead they continue to turn about half the distance normally required to make the signal switch to the next VU. This phenomenon might have lead to the increased turning accuracy in the study of Faugloire and Lejeune [56] for a similar condition. When using 8 VUs and each represents 45°, the deviation between indicated (D1) and actual (D2) directions lies between 0°-22.5°. Therefore, the expected mean deviation, i.e. the absolute error AE, is 11.25. Yet, in the action-perception coupling condition subjects reached a mean AE of 10.1°, 1.15° better than expected. Similar to their result, in our EB condition we received a significant better result than expected with AE = 4.9° instead of 5.6°. This is a similar performance increase of 10.2% and 12.5% respectively, possibly due to the same phenomenon.

The action-perception coupling might have such a strong effect that subjects in the MB condition pay less attention to adjusting the final position. The intuitive understanding of how much they have to turn to let the signal switch from one VU to the next one does not apply to the frontal ones, thus possibly annulling potential benefits of the MB direction distribution.

Another possible reason why AE was not lower for MB lies in the inaccuracy of the built-in digital compass of the smartphone that was used to present the tactile stimulation on the correct location and measured the orientation during the following movements. We were not able to find data about the accuracy of the Nexus 6, the smartphone model we used; however, errors of smartphone compasses have been found to be on average between 5° to 20° [17, 140] for previous generations of smartphones. The problem was controlled as much as possible by setting the initial direction for each trial to 0°. Additionally, when presenting the target direction D1 and measuring the actual direction D2, the error should be the same. The measured position stayed mostly within less than 1°when in a stable position. However, it is difficult to estimate the effect that inaccurate compass readings might have had on the turning accuracy. In conclusion, we have to reject the second hypothesis that suggested that the MB condition leads to better turning accuracy than the EB one.

The reaction time was indeed the same in both cases and with around 1.4s similar to the continuous vibration condition in the study of Cosgun et al. (see Figure 4 in [38]). Therefore, we can accept the third hypothesis, which states that the reaction time does not differ between the conditions.

5.7.2 Mental task

The results show that having additional mental load while performing the turning task does not have a negative effect on turning performance. Neither the accuracy of following the signal nor the reaction time decreases, even though participants had the subjective feeling they would perform worse in the combined task.

The effect on mental performance when simultaneously attending to the belt is small. When combining both tasks for the first time, the mental performance is slightly worse. However, when performing both tasks simultaneously for the second time, i.e. after about 90s of training,, the difference decreases and is not significant anymore. Klatzky et al. [88] found a similar result. In their study they measured the influence of mental load on a navigation task either guided by speech or by virtual sound. Virtual sound allowed for a similar action-perception coupling as in our experiment. Performance in this condition decreased during the first trial in which both tasks were combined. Already during the second combined trial performance resembled again that of the navigation task without the additional mental load.

Our results indicates that processing the tactile signal and turning as instructed does not require much effort and does not have a negative effect on mental tasks performed in parallel. Also, the average error rate does not change. Therefore, we can accept the fourth hypothesis, at least in part. No effect of the mental task on the turning task was found and performance in the mental task only slightly decreased. When performing both tasks for the second time, performance in the mental task already improved significantly. Hence, we assume training with a tactile belt can enhance parallel performance considerably[83].

5.8 Conclusion and outlook

In conclusion, to achieve high turning accuracy the MB was not superior to the EB. The mean AE of 4.9° is sufficient in most navigational purposes, supporting our design with 16 VUs. We did not improve turning accuracy by changing the angle distribution in the front (MB), possibly due to the effect of action-perception coupling. To avoid these problems, we will build a better compass directly in the tactile belt and use its data. We can then test similar settings again for specific tasks in which high turning accuracy is necessary.

When combining the turning task with the mental task, no negative effect on the turning task was detected. A negative effect on the mental task was only found during the first combination trial. The second time, no significant impairment was found anymore. In our experiments, subjects performed each task only for 3 minutes in total. Hence, we expect that prolonged training would lead to completely equal performance for the mental task with and without the parallel turning task. We will address this assumption in future work.

Overall, these results suggest that using a tactile belt for supporting navigation tasks can be helpful, especially when the user has to process other cognitive input in parallel. This is especially the case for blind people who use their ears to monitor their surroundings

Chapter 6

Summary and conclusion

In this chapter we summarize what we learned about communication between humans and technologies, and what needs to be considered to make it as enjoyable and effective as possible. In this thesis we consider two complementary approaches in HCI: I. How to help technology to better understand humans, and II. How to help humans to better understand technology. In the first part we investigated automatic emotion and personality recognition. In the second, we worked towards developing and advancing devices aiming at simplifying understanding and reducing stress for the user. Here we summarize our research and explore possible applications. Finally, we point out the contributions but also the limitations of this work.

6.1 Summary

The inner state of individuals has a great influence on their actions and their perception of the world [21]. Whether people feel stressed does not only depend on the external factors made responsible for it, but also on the current emotional state and their personality [97, 75]. However, people are often unaware of their inner state [172].

In this thesis at first we aspired to understand how to make the inner state accessible to technology without requiring to ask people directly but instead by employing physiological signals. So far, physiological signals have already been used by the field of affective computing and successful emotion recognition from physiological data has been achieved [193]. We followed this trend but additionally worked on Automatic Personality Recognition (APR). The younger research area of APR "targets the externalization process and is the task of inferring self-assessed personalities from machine detectable distal cues" [171]. The information analyzed for APR range from written texts, spoken words, and nonverbal cues, to behavior in the context of social media, mobile devices, and computer games [171]. While all these methods require collecting data about actual behavior of people, data collection for APR from physiological signals can be done passively and does not require the person to provide input explicitly. Automatic APR could allow various applications to automatically adapt information output to the needs of the individual. To help advance APR we conducted two successful studies on affective state and emotion recognition. Additionally, we published a database to support the respective community.

Later on we sought to help people to better understand information provided by technology. Since a large amount of information provided simultaneously can cause high cognitive load and stress, we worked on two wearable devices supposed to reduce stress in different contexts.

First we designed Spire, a wearable sensor that uses a physiological signal, namely breathing rate, to detect the psychological state of the person wearing it. In contrast to other physiological signals, breathing can be changed actively and therefore used to influence the psychological state. The correspondent application helps the user to increase awareness of the current mental state, and gives breathing advice to achieve a desired mental state.

Finally, we tested a tactile belt that transmits directional information via vibro-tactile signals. As current devices rarely use the tactile sense, this is an advantage. The information of the tactile belt does not have to compete for the same sensory attention, and the mental workload is reduced because information is perceived through different channels [181]. We compared two display modes for indicating directions. While we did not find a difference in the accuracy of following the direction indications between the two modes, we were able to show that our design is successful and leads to better results than any other tactile belt described in the literature [56, 74, 136]. Additionally, with a second experiment we showed that using the belt has only little impact on other cognitive tasks performed simultaneously, and training could reduce it even further.

6.2 Putting research into practice

A common challenge in the academic world is to accomplish the transition from tested research prototypes to applications serving users [49]. Often, researchers lack the knowledge and motivation necessary to bring a product to the market, while entrepreneurs, in contrast, lack the skills to conduct the research required to develop a useful consumer product. One main challenge is bridging this gap by bringing different people together or educating researchers in both skill sets. We approached this challenge by investigating how to transform our research results into applications that could solve problems in daily life. Further on, we describe scenarios in which the products resulting from our research are likely to be applied.

6.2.1 Emotion and Personality recognition

The interest in personality computing is increasing steadily (for a review see [171]). First efforts to use personality knowledge in applications are mostly concerned with applying perceived personality traits to the output of systems.

One study showed that adding personality to the synthetic voice of a GPS system increased its acceptance [126].

Our results are an initial step towards implicit APR via physiological signals. Automatically detecting the personality of users could increase the performance in recommender systems. One study [78] showed that users prefer a personality quiz over a rating system to teach a recommender system their preferences. They preferred the quiz both because of the consequent recommendations and the reduced initial effort. Another work [162] developed a recommender system that takes both the current affective state and the personality of the user into account. While the affective state was predicted using facial videos, for accessing personality traits, questionnaires were used. User comfort could potentially be increased even further in future implementations, when also personality recognition is done automatically.

The popularity of computer games increases and accordingly the correspondent research. Games have been used to understand personality, for example by analyzing user profiles [190] or by developing a game specifically aiming at personality detection [187]. By automatically detecting emotions and personality, the game environment could be adapted to enhance the experience for gamers accordingly.

6.2.2 Breathing sensor

The breathing sensor Spire has already been developed as a market ready product that is sold internationally. By measuring physiological signals, the sensor with the corresponding app is able to predict the mental state and provide feedback. People have used it successfully for decreasing tension, and improving their breathing patterns and calmness during work and daily life[1].

Patients affected by lung diseases such as asthma, might benefit from the breathing sensor. Future research will show whether it is possible to warn the

[1]https://www.spire.io/why-people-love-spire

patient when an asthma attack impends, or even reduce incidences by proposing preventive exercises.

6.2.3 The tactile belt

For different user groups there are different needs and opportunities that encourage using a tactile device to provide directional information.

A few research groups have already implemented tactile belts for various applications. McDaniel et al. [116] use vibration location and duration of 7 VUs to indicate the direction and distance of nearby people to blind users. Ferscha et al. [58] employ a tactile belt to indicate exits during an evacuation procedure. Simulations showed that such intuitive information can decrease panic growth and improve successful evacuation. Another example application for a vibrotactile belt is teaching how to dance. Rosenthal et al [139] successfully taught basic dance steps via tactile cues to demonstrate the usability of the designed belt.

For tourism tactile navigation support presents an attractive opportunity. Cities, tourist attractions, and hotels, could leverage on tactile navigation to improve their guests travel experience. For instance, bus travel agencies or hotels could lend tactile belts to their guests, enabling them to visit a town with less stress, as finding their way back to the starting location (travel bus, hotel, etc.) is guaranteed. For elderly people who feel insecure in navigation [66], this technology could increase the quality of travel experiences tremendously. In view of the demographic change the importance of supporting elderly, especially in navigation, will continue to increase [120].

A navigation device relying on the tactile sense, would especially benefit blind people, who need support for navigation and other daily activities that are more challenging without vision. Many supporting devices, such as common navigation devices, rely mostly on visual information displays. While various sensory substitution devices (SSDs) have been developed for the visually impaired (see section 5.2.2), most of them have never reached blind users outside of research laboratories [49]. Elli et al. identified various reasons partly responsible for the low adoption of supporting devices in the lives of blind people: (i) The affordability is often low due to small-scale manufacturing, (ii) research groups often built SSDs to prove concepts without considering comfort and aesthetics, and (iii) missing opportunities for purchase and consequent training make it difficult to access new technologies.

These pitfalls responsible for low adoption rates of new technologies should be avoided when implementing the tactile navigation belt. Therefore, members of the feelSpace research group formed the company feelSpace GmbH in Germany that develops a user-friendly tactile belt. The company prepares not only to sell the belt as a product, but also to offer training tailored to the specific needs of the respective user groups, such as the visually impaired.

6.2.4 Combining both approaches

Improving mutual understanding between humans and the technology they use will improve the adaptation of systems to users and therefore enhance their satisfaction. However, users individual differences are often ignored. Personalization is partly possible via customizable settings on devices and applications, but this process can be complicated and users might not necessarily be aware of the options best fitting their needs. Machine learning approaches have also been used. Liu et al. [106] addressed personalization by building an adaptive user interface enhanced by personalized learning of users' behavior. Yao [189] applied deep personalization to pedestrian navigation systems to support the population that has traumatic brain injuries.

Tailoring the information provided by applications to individual users is particularly important when these users are people with special needs. Moreover, adapting to individual differences and current states automatically, will not only benefit more challenged user groups, but also individual users. Changing information output settings based on the individuals inner state might help decreasing cognitive overload and consequent stress and frustration.

Emotion and personality play an important role for learning as well [55]. Previous works already predicted learner's emotions to personalize the tutor's feedback [32]. Fatahi et al. [54] additionally took learners' personality into account and incorporated human features in the interaction with the user. The system resulted in higher learning quality and more satisfied users. Future applications could increase their success by adapting the speed and the difficulty of the tasks not only according to the users performance but also the individual state.

6.3 Contributions

We investigated both ends of the communication between humans and technologies, and aimed at highlighting the importance of following an integrative approach. In the following we summarize our contributions.

6.3.1 Technology understanding humans

Publication on implicit user-centric personality recognition

To automatically learn about the inner state of technology users, we investigated how to detect emotions and personality traits from physiologically signals. While emotion detection has already been performed successfully, we are the first to show initial success in personality recognition. In addition, we used only commercially available wearable devices to grant ecological validity.

We published the results at the International Conference on Multimodal Interaction 2015 [174] (Chapter 2).

Publication on the inference of personality traits and affect schedule

Similarly to the previous study, in this we successfully performed both, personality and emotion recognition. Additionally, we achieved inference of affective schedule, i.e. general positive and negative affect, via physiological signals.

We published the study at International Conference on Automatic Face and Gesture Recognition 2015 [1] (Chapter 3).

Database made available

To support the Affective Computing community on enhancing and experimenting with emotion and personality recognition algorithms, we published our database[2] containing physiological data in response to emotional stimuli. This 36 participants database is not only one of the largest available, but also there is currently no similar database publicly available that contains additional personality data. The correspondent paper was submitted to the Journal Transactions of Affective Computing and is currently under review [173] (Chapter 2).

6.3.2 Humans understanding technology

User study with breathing sensor

In face of the growing amount of wearable devices and the information they can communicate to their users, it is important to make information processing easy in order to avoid information overload and stress. In the framework of the first internship of the author of this thesis we worked on designing the breathing sensor Spire and the correspondent smartphone app that intend to reduce stress. Through a set of iterations we found a comfortable way to show the notion of breathing to users. With a user study we demonstrated that the sensor helped people to reduce stress and be more aware of their mental state. Thousands of

[2] http://mhug.disi.unitn.it/index.php/datasets/ascertain/

sensors were sold by now and received great reviews on their ability to reduce stress, calm users down and help them to be more productive[3] (Chapter 4).

Investigation of applicability of a tactile belt for visually impaired people

Together with feelSpace we worked on advancing the tactile belt from a research prototype state towards a market ready version. While tactile belts have been tested in the lab with promising results (see section 5.3) especially for supporting navigation, so far no version has been released to the market that supports visually impaired or other users. Initially, we investigated the usability for navigational purposes especially for blind people (Section 5.2). With our results we won the first prize in a global student competition held at Virginia Tech[4] endowed with US$ 25.000.

Experiment to optimize tactile guidance

With an experiment we investigated the effectiveness of direction indication with a tactile belt equipped with 16 vibrating units. We measured accuracy in orienting towards an indicated direction with a turning task, and reached a mean error of only 4.9°, thus demonstrating the usefulness for navigation applications. When comparing two different tactile display modes no difference in performance or user preferences were found (Chapter 5).

Experiment to test the influence of mental load on tactile guidance

With the second experiment we demonstrated that indeed no significant interference was present between a cognitive task and the accuracy with which people could follow the tactile signal. Performance in the mental task declined slightly when both were required at the same time. However, after only 1.5 minutes of training, task performance was not significantly different anymore, as compared

[3] https://www.spire.io/why-people-love-spire
[4] http://vtkwglobal.com/winners/2014

to performing each task alone. These results suggest that even a short training time has already a significant effect on learning to integrate the tactile signals (Chapter 5).

Currently we prepare a paper with the results of the two experiments to submit it to the International Conference on Multimodal Interaction 2016.

The recently formed company feelSpace continues working on the tactile belt to put our research into practice. After her defense the author will become an employee of the company feelSpace.

6.4 Limitations and Future Work

In this section we explain the limitations of the research carried out throughout this thesis. We explain how the limitations impact our plans for future research and how these limitations could be overcome.

6.4.1 Emotion and personality recognition

We were one of the first to show that physiological signals in fact contain information that could be used to predict personality. However, in order to make useful applications able to automatically detect personality, more research is necessary to guarantee continuous accuracy of the recognition results, and to allow a consistent integration of emotion recognition results. We have not yet identified consistent rules on how personality mediates the emotional state under varying circumstances.

We are aware that the nature of personality measurements is quite complex. Personality traits are calculated from a compilation of various personal characteristics. It could be beneficial to test each personal characteristic on its own, and analyze which characteristics have the biggest effect on specific physiological signals. We will further try to understand, whether reaching higher recognition accuracies reliably and continuously is possible by improving the selection of sensors, choosing different analyzing methods or other emotional stimuli. We have made our database available so that other researchers could conduct such work as well.

In our recent studies, subjects were not equally distributed along the personality scales, possibly depriving us of valuable data from people with more prominent personality traits. We separated people in high and low-trait groups depending on the median. While this is a common practice in APR research [171], ranking people according to their personality traits might be psychologically more meaningful. For some personality traits the distribution along the scale was low and our subjects formed a more homogeneous group. Hence, the difference between the two binary groups was relatively small, and might have caused worse recognition results for the respective personality traits. To avoid this problem, for our next study we will consider preselecting our subjects according to their personality traits to include subjects with a bigger variance in personality.

6.4.2 Breathing sensor

While the breathing sensor was initially proven successful in our short-term user study for increasing awareness and reducing subjective stress, we did not conduct any long term study on the effects of using the breathing sensor on daily life. However, selling numbers seem to propose that the results of our work have a positive impact on customers' lives.

Unfortunately, because not all the raw data gathered by the sensor is available to users, future research on analyzing breathing patterns is limited.

6.4.3 Tactile belt

Comparing two response modes of the belt (EB and MB) both lead to successful accuracy of perceiving tactile cues that indicate directions. However, we could not show that one improves response accuracy over the other. We will investigate the necessary accuracy in real life applications to determine whether there are situations for which the achieved precision is not sufficient.

We could prove that cognitive load has no significant effect on the ability of using the tactile belt. The influence of the tactile belt on the cognitive task was small and already decreased after minimal training. In our future work we will evaluate whether training in fact facilitates complete integration of the tactile signal and eliminates any interference between the two parallel tasks.

In our study we controlled for having only two tasks, however, in real navigation situations more tasks might be present or movements could decrease the accuracy of perceiving the tactile signals. Therefore, we will carry out studies in more natural conditions with real navigation scenarios.

Bibliography

[1] Mojtaba Khomami Abadi, Juan Abdon Miranda Correa, Julia Wache, Heng Yang, Ioannis Patras, and Nicu Sebe. Inference of personality traits and affect schedule by analysis of spontaneous reactions to affective videos. In *11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, 2015.

[2] Mojtaba Khomami Abadi, Seyed Mostafa Kia, Ramanathan Subramanian, Paolo Avesani, and Nicu Sebe. User-centric Affective Video Tagging from MEG and Peripheral Physiological Responses. In *Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 582–587, 2013.

[3] Mojtaba Khomami Abadi, Ramanathan Subramanian, Seyed Mostafa Kia, Paolo Avesani, Ioannis Patras, and Nicu Sebe. Decaf: Meg-based multimodal database for decoding affective physiological responses. *IEEE Transaction on Affective Computing*, 2015.

[4] F. Q. Al-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan. Respiration rate monitoring methods: a review. *Pediatric pulmonology*, 46(6):523–9, jun 2011.

[5] Xavier Alameda-Pineda, Jacopo Staiano, Ramanathan Subramanian, Ligia M. Batrinca, Elisa Ricci, Bruno Lepri, Oswald Lanz, Nicu Sebe, and Senior Member. SALSA: A Novel Dataset for Multimodal Group Behavior Analysis. *arXiv*, (i):1–14, 2015.

[6] Anton Aluja, Óscar García, and Luís F. García. Relationships among extraversion, openness to experience, and sensation seeking. *Personality and Individual Differences*, 35(3):671–680, 2003.

[7] S. Argamon, S. Dhawle, M. Koppel, and James W. Pennbaker. Lexical predictors of personality type. *Interface and the Classification Society of North America*, jan 2005.

[8] Amna Asif, Wilko Heuten, and Susanne Boll. Exploring Distance Encodings with a Tactile Display to Convey Turn by Turn Information in Automobiles. *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*, 2010.

[9] Bruce A. Austin. Motivations for movie attendance. *Communication Quarterly*, 34(2):115–126, 1986.

[10] Paul Bach-y Rita. Vision substitution by tactile image projection. *Nature*, 21:963–964, 1969.

[11] Paul Bach-y Rita. Tactile sensory substitution studies. *Annals of the New York Academy of Sciences*, 1013:83–91, 2004.

[12] Paul Bach-y Rita, K A Kaczmarek, M E Tyler, and J Garcia-Lara. Form perception with a 49-point electrotactile stimulus array on the tongue: a technical note. *Journal of rehabilitation research and development*, 35(4):427–430, 1998.

[13] Paul Bach-y Rita and Stephen W. Kercel. Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7(12):541–546, 2003.

[14] A. Batliner, C. Hacker, S. Steidl, and E. Nöth. 'You Stupid Tin Box'-Children Interacting with the AIBO Robot: A Cross-linguistic Emotional Speech Corpus. *Lrec*, pages 171–174, 2004.

[15] Anton Batliner, K. Fischer, R. Huber, J. Spilker, and E. Nöth. How to find trouble in communication. *Speech Communication*, 40(1-2):117–143, 2003.

[16] Yoav Benjamini and Yosef Hochberg. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. 57(1):289 – 300, 1995.

[17] Jeffrey R. Blum, D Greencorn, and Jeremy R. Cooperstock. Smartphone sensor reliability for augmented reality applications. *Mobile and ubiquitous systems: computing, networking, and services*, pages 127–138, 2013.

[18] Susanne Boll, Amna Asif, and Wilko Heuten. Feel your route: a tactile display for car navigation. *IEEE Pervasive Computing*, 10(3):35–42, 2011.

[19] Erik Borg, Jerker Rönnberg, and Lennert Neovius. Vibratory-coded directional analysis: Evaluation of a three-microphone/four-vibrator DSP sysetem. *Journal of rehabilitation research and development*, 38(2):257–263, 2001.

[20] Robert F Bornstein. Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological bulletin*, 106(2):265, 1989.

[21] B P Bradley and K Mogg. Mood and personality in recall of positive and negative information. *Behaviour research and therapy*, 32(1):137–41, jan 1994.

[22] Anne-Marie Brouwer, Martin G. van Schaik, J. E. Korteling, Jan B. F. van Erp, and Alexander Toet. Neuroticism, Extraversion, Conscientiousness and Stress: Physiological Correlates. *IEEE Transactions on Affective Computing*, 6(2):109–117, apr 2015.

[23] Gavin Brown, Adam Pocock, Ming-Jie Zhao, and Mikel Lujan. Conditional Likelihood Maximisation: A Unifying Framework for Mutual Information Feature Selection. *Journal of Machine Learning Research*, 13:27–66, 2012.

[24] Lindsay Brown, Bernard Grundlehner, and Julien Penders. Towards wireless emotional valence detection from EEG. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2188–2191. IEEE, aug 2011.

[25] Richard P. Brown and Patricia L. Gerbarg. Sudarshan Kriya Yogic Breathing in the Treatment of Stress, Anxiety, and Depression: Part II Clinical Applications and Guidelines. *The Journal of Alternative and Complementary Medicine*, 11(4):711–717, 2005.

[26] Richard P. Brown and Patricia L. Gerbarg. Sudarshan Kriya Yogic Breathing in the Treatment of Stress, Anxiety, and Depression: Part INeurophysiologic Model. *The Journal of Alternative and Complementary Medicine*, 11(1):189–201, feb 2005.

[27] Claudia Chloe Brumbaugh, Ravi Kothuri, Carl Marci, Caleb Siefert, and Donald D. Pfaff. Physiological correlates of the Big 5: Autonomic responses to video presentations. *Applied Psychophysiology Biofeedback*, 38(4):293–301, 2013.

[28] Volkert Buchmann, Mark Billinghurst, and Andy Cockburn. Directional interfaces for wearable augmented reality. In *Proceedings of the 9th ACM SIGCHI New Zealand Chapter's International Conference on Human-Computer Interaction Design Centered HCI - CHINZ '08*, pages 47–54, 2008.

[29] Xavier P. Burgos-Artizzu, Pietro Perona, and Piotr Dollar. Robust Face Landmark Estimation under Occlusion. In *2013 IEEE International Conference on Computer Vision*, pages 1513–1520. IEEE, dec 2013.

[30] George Caridakis, Lori Malatesta, Loic Kessous, and Noam Amir. Modeling naturalistic affective states via facial and vo- cal expressions recognition. 2004.

[31] A. Cassinelli, E. Sampaio, S. B. Joffily, H. Lima, and B. Gusmão. Do blind people move more confidently with the Tactile Radar? *Technology and Disability*, 26(2-3):161–170, 2014.

[32] Soumaya Chaffar, Gerardo Cepeda, and Claude Frasson. Predicting the Learner's Emotional Reaction towards the Tutor's Intervention. In *Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, pages 639–641. IEEE, jul 2007.

[33] Seungmoon Choi and Katherine J. Kuchenbecker. Vibrotactile Display: Perception, Technology, and Applications. *Proceedings of the IEEE*, 101(9):2093–2104, sep 2013.

[34] Roger W. Cholewiak, J Christopher Brill, and Anja Schwab. Vibrotactile localization on the abdomen: Effects of place and space. *Perception & Psychophysics*, 66(6):970–987, aug 2004.

[35] Roger W. Cholewiak and C. McGrath. Vibrotactile Targeting in Multimodal Systems: Accuracy and Interaction. In *14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pages 413–420. IEEE, 2006.

[36] Michael E. Clark and Richard Hirschman. Effects of paced respiration on anxiety reduction in a clinical population. *Biofeedback and self-regulation*, 15(3):273–84, sep 1990.

[37] Ansgar Conrad, Anett Müller, Sigrun Doberenz, Sunyoung Kim, Alicia E. Meuret, Eileen Wollburg, and Walton T. Roth. Psychophysiological effects of breathing instructions for stress management. *Applied psychophysiology and biofeedback*, 32(2):89–98, jun 2007.

[38] Akansel Cosgun, Emrah Akin Sisbot, and Henrik I. Christensen. Evaluation of rotational and directional vibration patterns on a tactile belt for guiding visually impaired people. In *2014 IEEE Haptics Symposium (HAPTICS)*, pages 367–370. IEEE, feb 2014.

[39] Paul T Costa and Robert R McCrae. Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, 38(4):668–678, 1980.

[40] Paul T Jr. Costa and Robert R McCrae. *NEO-PI-R professional manual: Revised NEO personality and NEO Five-Factor Inventory (NEO-FFI)*, volume 4. Psychological Assessment Resources, 1992.

[41] Dimitrios Dakopoulos and N.G. Bourbakis. Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1):25–35, jan 2010.

[42] Navneet Dalal and Bill Triggs. Histograms of Oriented Gradients for Human Detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, volume 1, pages 886–893, 2005.

[43] Bradley M. Davis. Effects of Tactical Navigation Display Modality on Navigation Performance, Situation Awareness, and Mental Workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(17):2089–2093, 2006.

[44] Vilfredo De Pascalis, Enrica Strippoli, Patrizia Riccardi, and Fabiola Vergari. Personality, event-related potential (ERP) and heart rate (HR) in emotional word processing. *Personality and Individual Differences*, 36(4):873–891, mar 2004.

[45] M. C. Dorneich, P. M. Ververs, S. D. Whitlow, and S. Mathan. Evaluation of a Tactile Navigation Cueing System and Real-Time Assessment of Cognitive State. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(24):2600–2604, oct 2006.

[46] Ellen Douglas-Cowie, Roddy Cowie, Ian Sneddon, Cate Cox, Orla Lowry, Margaret Mcrorie, Jean-claude Martin, Laurence Devillers, Sarkis Abrilian, Anton Batliner, Noam Amir, and Kostas Karpouzis. The HUMAINE Database : Addressing the Collection and Annotation of Naturalistic and Induced Emotional Data. In *Affective computing and intelligent interaction*, pages 488–500. 2007.

[47] Paul Ekman and Wallace v. Friesen. Constants across cultures in the face and emotion, 1971.

[48] Moataz El Ayadi, Mohamed S. Kamel, and Fakhri Karray. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3):572–587, mar 2011.

[49] Giulia V. Elli, Stefania Benetti, and Olivier Collignon. Is There a Future for Sensory Substitution Outside Academic Laboratories? *Multisensory Research*, 27(5-6):271–291, nov 2014.

[50] Linda R Elliott, Michael D Coovert, and Elizabeth S Redden. Overview of meta-analyses investigating vibrotactile versus visual display options. In *Human-Computer Interaction. Novel Interaction Methods and Techniques*, pages 435–443. Springer, 2009.

[51] Linda R. Elliott, Jan B. F. van Erp, Elizabeth S. Redden, and Maaike Duistermaat. Field-Based Validation of a Tactile Navigation Device. *IEEE Transactions on Haptics*, 3(2):78–87, apr 2010.

[52] Hans J. Eysenck. *Dimensions of Personality*. Transaction Publishers, 5 edition, 1950.

[53] Beat Fasel and Juergen Luettin. Automatic facial expression analysis: a survey. *Pattern Recognition*, 36(1):259–275, jan 2003.

[54] S. Fatahi, M. Kazemifard, and N. Ghasem-Aghaee. Design and Implementation of an Intelligent Educational Model Based on Personality and Learner's Emotion. *Advances in Electrical Engineering and Computational Science*, pages 423–434, 2010.

[55] Somayeh Fatahi and Nasser Ghasem-Aghaee. Design and Implementation of an Intelligent Educational Model Based on Personality and Learner's Emotion. *Advances in Electrical Engineering and Computational Science*, 7(3):423–434, 2010.

[56] Elise Faugloire and Laure Lejeune. Evaluation of heading performance with vibrotactile guidance: The benefits of informationmovement coupling compared with spatial language. *Journal of Experimental Psychology: Applied*, 20(4):397–410, 2014.

[57] Alois Ferscha, Bernadette Emsenhuber, Andreas Riener, Clemens Holzmann, Manfred Hechinger, Dominik Hochreiter, Marquart Franz, and Andreas Zeidler. Vibro-Tactile Space-Awareness. In *International Conference on Ubiquitous Computing*, 2008.

[58] Alois Ferscha and Kashif Zia. LifeBelt: Silent Directional Guidance for Crowd Evacuation. In *International Symposium on Wearable Computers*, number August, pages 19–26. IEEE, sep 2009.

[59] German Flores, Sri Kurniawan, Roberto Manduchi, Eric Martinson, Lourdes M. Morales, and Emrah Akin Sisbot. Vibrotactile Guidance for Wayfinding of Blind Walkers. *IEEE Transactions on Haptics*, 8(3):306–317, jul 2015.

[60] N. Fragopanagos and J.G. Taylor. Emotion recognition in humancomputer interaction. *Neural Networks*, 18(4):389–405, may 2005.

[61] Barbara L. Fredrickson, Roberta A. Mancuso, Christine Branigan, and Michele M. Tugade. The Undoing Effect of Positive Emotions. *Motivation and Emotion*, 24(4):237–258, 2000.

[62] Brenda O Gilbert. Physiological and Nonverbal Correlations of Extraversion, Neuroticism, and Psychoticism during Active and Passive Coping. *Personality and individual differences*, 12(12):1325–1331, 1991.

[63] Nicholas A. Giudice and Gordon E. Legge. Blind Navigation and the Role of Technology. In *The Engineering Handbook of Smart Technology for Aging, Disability, and Independence*, pages 479–500. John Wiley & Sons, Inc., jan 2008.

[64] Louis H Goldish and Harry E Taylor. The Optacon: A Valuable Device for Blind Persons. *New Outlook for the Blind*, 68(2):49–56, 1974.

[65] Reginald G. Golledge, James R. Marston, Jack M. Loomis, and Roberta L. Klatzky. Stated preferences for components of a personal guidance system for nonvisual navigation. *Journal of Visual Impairment & Blindness*, 98(3):135–147, 2004.

[66] Lawrence E. M. Grierson, John S. Zelek, and Heather Carnahan. The Application of a Tactile Way-finding Belt to Facilitate Navigation in Older Persons. *Ageing International*, 34(4):203–215, dec 2009.

[67] Lawrence E. M. Grierson, John S. Zelek, Isabel Lam, Sandra E. Black, and Heather Carnahan. Application of a Tactile Way-Finding Device to Facilitate Navigation in Persons With Dementia. *Assistive Technology*, 23(2):108–115, jun 2011.

[68] James J. Gross and Robert W. Levenson. Emotion elicitation using films. *Cognition & Emotion*, 9(1):87–108, jan 1995.

[69] E. Grossman, A. Grossman, M. H. Schein, R. Zimlichman, and B. Gavish. Breathing-control lowers blood pressure. *Journal of human hypertension*, 15(4):263–9, apr 2001.

[70] Sunjai Gupta and John Nicholson. Simple Visual Reaction Time, Personality and strength of the Nervous System: A signal-detection theory approach. *Personality and Individual Differences*, 6(4):461–469, 1985.

[71] L Haber, R N Haber, S Penningroth, K Novak, and H Radgowski. Comparison of nine methods of indicating the direction to objects: data from blind adults. *Perception*, 22(1):35–47, 1993.

[72] John E. Hall and Arthur C. Guyton. Behavioral and Motivational Mechanisms of the BrainThe Limbic System and the Hypothalamus. In *Guyton and Hall Textbook of Medical Physiology*, chapter 59, pages 711–718. 12 edition, 2011.

[73] Karla R. Hamlen and Thomas J Shuell. The Effects of Familiarity and Audiovisual Stimuli on Preference for Classical Music. *Bulletin of the Council for Research in Music Education*, 168:21–34, 2006.

[74] Wilko Heuten, Niels Henze, Susanne Boll, and Martin Pielot. Tactile Wayfinder: A Non-Visual Support System for Wayfinding. In *Proceedings of the 5th Nordic conference on Human-computer interaction building bridges - NordiCHI '08*, page 172, 2008.

[75] Peter Hills and Michael Argyle. Emotional stability as a major dimension of happiness. *Personality and Individual Differences*, 31(8):1357–1364, dec 2001.

[76] Cristy Ho, Nick Reed, and Charles Spence. Assessing the effectiveness of ”intuitive” vibrotactile warning signals in preventing front-to-rear-end collisions in a driving simulator. *Accident Analysis and Prevention*, 38(5):988–996, 2006.

[77] Cristy Ho, Hong Z. Tan, and Charles Spence. Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(6):397–412, 2005.

[78] Rong Hu and Pearl Pu. A Comparative User Study on Rating vs. Personality Quiz based Preference Elicitation Methods. In *Proceedings of the 2009 International Conference on Intelligent User Interfaces*, pages 367–371, 2009.

[79] Lise A. Johnson and Charles M. Higgins. A navigation aid for the blind using tactile-visual sensory substitution. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pages 6289–6292, 2006.

[80] Hideo Joho, Jacopo Staiano, Nicu Sebe, and Joemon M. Jose. Looking at the Viewer: Analysing Facial Activities to Detect Personal Highlights of Multimedia Contents. *Multimedia Tools and Applications*, 51(2):505–523, 2011.

[81] Lynette A. Jones and Nadine B. Sarter. Tactile Displays: Guidance for Their Design and Application. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(1):90–111, 2008.

[82] Kurt A. Kaczmarek. Sensory augmentation and substitution. *CRC handbook of biomedical engineering*, pages 2100–2109, 1995.

[83] Silke M. Kärcher, Sandra Fenzlaff, Daniela Hartmann, Saskia K. Nagel, and Peter König. Sensory Augmentation for the Blind. *Frontiers in Human Neuroscience*, 6(March):1–15, 2012.

[84] Kai Kaspar, Sabine König, Jessika Schwandt, and Peter König. The experience of new sensorimotor contingencies by sensory augmentation. *Consciousness and Cognition*, 28(1):47–63, 2014.

[85] Elizabeth G. Kehoe, John M. Toomey, Joshua H. Balsters, and Arun L W Bokde. Personality modulates the effects of emotional arousal and valence on brain activation. *Social Cognitive and Affective Neuroscience*, 7(7):858–870, oct 2012.

[86] Jonghwa Kim and Elisabeth André. Emotion recognition based on physiological changes in music listening. *IEEE transactions on pattern analysis and machine intelligence*, 30(12):2067–83, dec 2008.

[87] Kyung Hwan Kim, Seok Won Bang, and Sang Ryong Kim. Emotion recognition system using short-term monitoring of physiological signals. *Medical & biological engineering & computing*, 42(3):419–27, may 2004.

[88] Roberta L Klatzky, James R Marston, Nicholas A. Giudice, Reginald G Golledge, and Jack M Loomis. Cognitive load of navigating without vision when guided by virtual sound versus spatial language. *Journal of Experimental Psychology: Applied*, 12(4):223–232, 2006.

[89] Sander Koelstra, Christian Mühl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and

Ioannis Patras. DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Transactions on Affective Computing*, 3(1):18–31, 2012.

[90] Sander Koelstra and Ioannis Patras. Fusion of facial expressions and EEG for implicit affective tagging. *Image and Vision Computing*, 31(2):164–174, feb 2013.

[91] Sander Koelstra, Ashkan Yazdani, Mohammad Soleymani, and M Christian. Single Trial Classification of EEG and Peripheral Physiological Signals for Recognition of Emotions Induced by Music Videos. *Brain informatics*, pages 1–12, 2010.

[92] Andrew J. Kolarik, Silvia Cirstea, Shahina Pardhan, and Brian C.J. Moore. A summary of research investigating echolocation abilities of blind and sighted humans. *Hearing Research*, 310:60–68, apr 2014.

[93] Gerbert Kraaykamp and Koen Van Eijck. Personality, media preferences, and cultural participation. *Personality and Individual Differences*, 38(7):1675–1688, may 2005.

[94] O. Lahav, H. Gedalevitz, S. Battersby, D. Brown, L. Evett, and P. Merritt. Using Wii technology to explore real spaces via virtual environments for people who are blind. *Journal of Assistive Technologies*, 8(3):150–160, sep 2014.

[95] Peter J Lang, Margaret M. Bradley, and B N Cuthbert. International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual. Technical report, University of Florida, 2008.

[96] Peter J. Lang, Mark K. Greenwald, Margaret M. Bradley, and Alfons O. Hamm. Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30(3):261–73, may 1993.

[97] Randy J. Larsen and Timothy Ketelaar. Extraversion, Neuroticism and Susceptibility to Positive and Negative Mood Induction Procedures. *Personality and Individual Differences*, I(12):1221–1228, 1989.

[98] Chul Min Lee and Shrikanth S Narayanan. Toward detecting emotions in spoken dialogs. *Speech and Audio Processing*, 13(2):293–303, 2005.

[99] Charles Lenay, Olivier Gapenne, and John Stewart. The Constitution of Spatiality in Relation to the Lived Body : a Study based on Prosthetic Perception. *Conference on the Emergence and Development of Embodied Cognition*, 2001.

[100] Bruno Lepri, Ramanathan Subramanian, Kyriaki Kalimeri, Jacopo Staiano, Fabio Pianesi, and Nicu Sebe. Connecting Meeting Behavior with Extraversion – A Systematic Study. *IEEE Transactions on Affective Computing*, 3(4):443–455, jan 2012.

[101] David M. Levy, Jacob O. Wobbrock, Alfred W. Kaszniak, and Marilyn Ostergren. The Effects of Mindfulness Meditation Training on Multitasking in a High-Stress Information Environment. In *Graphics Interface 2012*, pages 45–52, Toronto, Canada, 2012.

[102] Ronald Ley. The Modification of Breathing Behavior: Pavlovian and Operant Control in Emotion and Cognition. *Behavior Modification*, 23(3):441–479, jul 1999.

[103] Yueqing Li, Woo Ram Jeon, and Chang S Nam. Navigation by vibration: Effects of vibrotactile feedback on a navigation task. *International Journal of Industrial Ergonomics*, 46(0):76–84, 2015.

[104] Tanja Lischetzke and Michael Eid. Why extraverts are happier than introverts: the role of mood regulation. *Journal of personality*, 74(4):1127–61, aug 2006.

[105] Christine L. Lisetti and Fatma Nasoz. Using Noninvasive Wearable Computers to Recognize Human Emotions from Physiological Signals. *EURASIP Journal on Advances in Signal Processing*, 11(11):1672–1687, 2004.

[106] Jiming Liu, Chi Kuen Wong, and Ka Keung Hui. An adaptive user interface based on personalized learning. *IEEE Intelligent Systems*, (2):52–57, 2003.

[107] Jack M. Loomis and Susan J. Lederman. Tactual Perception. In *Handbook of perception and human performances 2*, volume 2. jun 1986.

[108] Jack M Loomis, James R Marston, Reginald G Golledge, and Roberta L Klatzky. Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance. *Journal of visual impairment & blindness*, 99(4):219–232, 2005.

[109] Patrick Lucey, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, number July, pages 94–101. IEEE, jun 2010.

[110] Shachar Maidenbaum, Sami Abboud, and Amir Amedi. Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation. *Neuroscience & Biobehavioral Reviews*, 41:3–15, apr 2014.

[111] Shachar Maidenbaum, Shlomi Hanassy, Sami Abboud, Galit Buchs, Daniel Robert Chebat, Shelly Levy-Tzedek, and Amir Amedi. The 'Eye-Cane', a new electronic travel aid for the blind: Technology, behavior &

swift learning. *Restorative Neurology and Neuroscience*, 32(6):813–824, 2014.

[112] Francois Mairesse, Marilyn A. Walker, Matthias R. Mehl, and Roger K. Moore. Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. *Journal of Artificial Intelligence Research*, 30:457–500, 2007.

[113] P. Manivannan and S. Bharathiraja. Qi Open Wireless Charging Standard A Wireless Technology for the Future. *International journal of engineering and computer science*, 2(3):573–579, 2013.

[114] James R Marston, Jack M Loomis, Roberta L Klatzky, and Reginald G Golledge. Nonvisual route following with guidance from a simple haptic or auditory display. *Journal of Visual Impairment & Blindness*, 101(4):203–211, 2007.

[115] Robert R McCrae and O P John. An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215, jun 1992.

[116] Troy McDaniel, Sreekar Krishna, Vineeth Balasubramanian, Dirk Colbry, and Sethuraman Panchanathan. Using a haptic belt to convey non-verbal communication cues during social interactions to individuals who are blind. *HAVE 2008 - IEEE International Workshop on Haptic Audio Visual Environments and Games Proceedings*, (October):13–18, 2008.

[117] Jane A McElroy and Michael C Hosokawa. Take a Deep Breath : A Pilot Study Demonstrating a Significant Reduction in Blood Pressure with 15 Minute Daily Pranayama Breathing. *Annals of Behavioral Science and Medical Education*, 18(2):15–18, 2012.

[118] Lotfi B. Merabet, Erin C. Connors, Mark A. Halko, and Jaime Sánchez. Teaching the Blind to Find Their Way by Playing Video Games. *PLoS ONE*, 7(9):e44958, sep 2012.

[119] Ivan Mervielde, Filip De Fruyt, and Slawomir Jarmuz. Linking openness and intellect in childhood and adulthood. *Parental descriptions of child personality: Developmental antecedents of the Big Five*, pages 105–126, 1998.

[120] Scott D. Moffat. Aging and spatial navigation: What do we know and where do we go? *Neuropsychology Review*, 19(4):478–489, 2009.

[121] Neema Moraveji, Takehiro Hagiwara, and Athman Adiseshan. Breath-Tray: Influencing Self- Regulation without Cognitive Deficit. In *CHI 2012*, pages 2405–2410, Austin, Texas, 2012.

[122] Neema Moraveji, Ben Olson, Truc Nguyen, Mahmoud Saadat, Yaser Khalighi, Roy Pea, and Jeffrey Heer. Peripheral Paced Respiration : Influencing User Physiology during Information Work. In *UIST'11*, pages 423–427, Santa Barbara, 2011.

[123] Neema Moraveji and Charlton Soesanto. Towards stress-less user interfaces: 10 Design Heuristics Based on the Psychophysiology of Stress. *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*, page 1643, 2012.

[124] Saskia K. Nagel, Christine Carl, Tobias Kringe, Robert Märtin, and Peter König. Beyond sensory substitution–learning the sixth sense. *Journal of neural engineering*, 2(4):R13–26, dec 2005.

[125] Fatma Nasoz, Kaye Alvarez, Christine L. Lisetti, and Neal Finkelstein. Emotion Recognition from Physiological Signals for Presence Technolo-

gies. *International Journal of Cognition, Technology, and Work - Special Issue on Presence*, 6(1), 2003.

[126] Clifford Ivar Nass and Scott Brave. *Wired for speech: How voice activates and advances the human-computer relationship*. MIT press Cambridge, 2005.

[127] Weiting Ng. Clarifying the relation between neuroticism and positive emotions. *Personality and Individual Differences*, 47(1):69–72, jul 2009.

[128] Jakob Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems - CHI '94*, volume 3, pages 413–414, New York, New York, USA, 1994. ACM Press.

[129] Daniel Olguín Olguín, Benjamin N. Waber, Taemie Kim, Akshay Mohan, Koji Ara, and Alex Pentland. Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 39(1):43–55, 2009.

[130] J. Kevin O'Regan and Alva Noë. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(05):939–973, oct 2001.

[131] Maja Pantic and L.J.M. Rothkrantz. Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, 91(9):1370–1390, sep 2003.

[132] Maja Pantic, Michel Valstar, Ron Rademaker, and Ludo Maat. Web-Based Database for Facial Expression Analysis. In *2005 IEEE International Conference on Multimedia and Expo*, volume 2005, pages 317–321. IEEE, 2005.

[133] Marco Perugini and Lisa Di Blas. Analyzing personality-related adjectives from an eticemic perspective: The Big Five Marker Scales (BFMS) and the Italian AB5C taxonomy. *Big Five Assessment*, pages 281–304, 2002.

[134] Elizabeth A. Phelps. Emotion and cognition: insights from studies of the human amygdala. *Annual review of psychology*, 57(Miller 2003):27–53, jan 2006.

[135] Martin Pielot and Susanne Boll. Tactile Wayfinder: Comparison of tactile waypoint navigation with commercial pedestrian navigation systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6030 LNCS:76–93, 2010.

[136] Martin Pielot, Niels Henze, Wilko Heuten, and Susanne Boll. Evaluation of continuous direction encoding with tactile belts. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 5270 LNCS, pages 1–10, 2008.

[137] Oudeyer Pierre-Yves. The production and recognition of emotions in speech: features and algorithms. *International Journal of Human-Computer Studies*, 59(1):157–183, 2003.

[138] Maurice Ptito, Solvej M Moesgaard, Albert Gjedde, and Ron Kupers. Cross-modal plasticity revealed by electrotactile stimulation of the tongue in the congenitally blind. *Brain*, 128(3):606–614, 2005.

[139] Jacob Rosenthal, Nathan Edwards, Daniel Villanueva, Sreekar Krishna, Troy McDaniel, and Sethuraman Panchanathan. Design, Implementation, and Case Study of a Pragmatic Vibrotactile Belt. *IEEE Transactions on Instrumentation and Measurement*, 60(1):114–125, jan 2011.

[140] Laura Ruotsalainen, Heidi Kuusniemi, and Ruizhi Chen. Heading change detection for indoor navigation with a Smartphone camera. *2011 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2011*, (September):21–23, 2011.

[141] James A. Russell. A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[142] James A. Russell and Albert Mehrabian. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3):273–294, 1977.

[143] Stefanie Schmidt and Wolfgang G. Stock. Collective Indexing of Emotions in Images. A Study in Emotional Information Retrieval. *Journal of the American Society for Information Science and Technology*, 60(5):863–876, 2009.

[144] Nicu Sebe, Ira Cohen, Theo Gevers, and T.S. Huang. Emotion Recognition Based on Joint Visual and Audio Cues. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 1, pages 1136–1139. IEEE, 2006.

[145] Nicu Sebe, Ira Cohen, and Thomas S. Huang. Multimodal emotion recognition. In *Handbook of Pattern Recognition and Computer Vision 4*, chapter 1, pages 387–419. 2005.

[146] Jae Woong Shim and Bryant Paul. Effects of Personality Types on the Use of Television Genre. *Journal of Broadcasting & Electronic Media*, 51(2):287–304, jul 2007.

[147] Ian Sinclair, Jim Carter, Sebastian Kassner, Jan B. F. Van Erp, Gerhard Weber, Linda R. Elliott, and Ian Andrew. Towards a standard on evaluation of tactile/haptic interactions. *International Conference on Haptics:*

Perception, Devices, Mobility, and Communication, EuroHaptics 2012, 7282 LNCS:528–539—-, 2012.

[148] Rajita Sinha. Multivariate Response Patterning of Fear and Anger. *Cognition & Emotion*, 10(2):173–198, mar 1996.

[149] Mohammad Soleymani, Jeroen Lichtenauer, Thierry Pun, and Maja Pantic. A Multimodal Database for Affect Recognition and Implicit Tagging. *IEEE Transactions on Affective Computing*, 3(1):42–55, jan 2012.

[150] Mayuree Srikulwong and Eamonn O'Neill. A comparative study of tactile representation techniques for landmarks on a wearable device. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, pages 2029–2038, New York, New York, USA, 2011. ACM Press.

[151] Mayuree Srikulwong and Eamonn O'Neill. Wearable tactile display of directions for pedestrian navigation: Comparative lab and field evaluations. In *2013 World Haptics Conference (WHC)*, volume 61, pages 503–508. IEEE, apr 2013.

[152] Ruchir Srivastava, Jiashi Feng, Sujoy Roy, Terence Sim, and Shuicheng Yan. Don't Ask Me What I'm Like, Just Watch and Listen. In *Proceedings of the 20th ACM international conference on Multimedia. ACM, 2012.*, pages 329–338, 2012.

[153] Jacopo Staiano, Bruno Lepri, Nadav Aharony, Fabio Pianesi, Nicu Sebe, and Alex Pentland. Friends Don't Lie: Inferring Personality Traits from Social Network Structure. In *ACM Conference on Ubiquitous Computing*, pages 321–330, 2012.

[154] Georg Stenberg. Personality and the EEG Arousal and Emotional Arousability. *Personality and Individual Differences*, 13(10):1097–1113, 1992.

[155] Con Stough, Catherine Donaldson, Belinda Scarlata, and Joseph Ciorciari. Psychophysiological correlates of the NEO PI-R Openness, Agreeableness and Conscientiousness: preliminary results. *International Journal of Psychophysiology*, 41(1):87–91, may 2001.

[156] Alexander Strehl and Joydeep Ghosh. Cluster Ensembles A Knowledge Reuse Framework for Combining Multiple Partitions. *Journal of Machine Learning Research*, 3:583–617, 2002.

[157] Ella Striem-Amit, Miriam Guendelman, and Amir Amedi. 'Visual' acuity of the congenitally blind using visual-to-auditory sensory substitution. *PloS one*, 7(3), 2012.

[158] C. Studholme, D.L.G. Hill, and D.J. Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32(1):71–86, jan 1999.

[159] Ramanathan Subramanian, Yan Yan, Jacopo Staiano, Lanz, Oswald, and Nicu Sebe. On the relationship between head pose, social attention and personality prediction for unstructured and dynamic group interactions. In *Int'l Conference on Multimodal Interaction*, pages 3–10, 2013.

[160] Hong Z Tan, Robert Gray, J Jay Young, and Ryan Traylor. A Haptic Back Display for Attentional and Directional Cueing. *Haptics-e*, 3(1):1–20, 2003.

[161] Hui Tang and David Beebe. An oral tactile interface for two-way communication. In *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology. Proceedings (Cat. No.00EX451)*, volume 14, pages 639–643. IEEE, 2000.

[162] Marko Tkalčič, Jurij Franc Tasič, and Andrej Košir. Emotive and Personality Parameters in Multimedia Recommender Systems. *Affective Com-*

puting and Intelligent Interaction PROCEEDINGS OF THE DOCTORAL CONSORTIUM 2009, page 33, 2009.

[163] Serdar Tok, Mehmet Koyuncu, Seda Dural, and Fatih Catikkas. Evaluation of International Affective Picture System (IAPS) ratings in an athlete population and its relations to personality. *Personality and Individual Differences*, 49(5):461–466, oct 2010.

[164] Dzmitry Tsetserukou, Junichi Sugiyama, and Jun Miura. Belt tactile interface for communication with mobile robot allowing intelligent obstacle detection. *2011 IEEE World Haptics Conference, WHC 2011*, pages 113–118, 2011.

[165] Koji Tsukada and M Yasumura. *ActiveBelt: Belt-Type Wearable Tactile Display for Directional Navigation*, volume 3205 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[166] Jan B. F. van Erp. Presenting directions with a vibrotactile torso display. *Ergonomics*, 48(3):302–313, 2005.

[167] Jan B. F. van Erp and Hendrik A. H. C. Van Veen. Vibrotactile in-vehicle navigation system. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(4-5):247–256, 2004.

[168] Jan B. F. van Erp, Hendrik A. H. C. van Veen, Chris Jansen, and Trevor Dobbins. Waypoint navigation with a vibrotactile waist belt. *ACM Transactions on Applied Perception*, 2(2):106–117, 2005.

[169] Jan B. F. van Erp, Hans Veltman, Hendrik A. H. C. van Veen, and A. B. Oving. Tactile torso display as countermeasure to reduce night vision goggles induced drift. In *NATO RTO HFM panel symposium Spatial disorientation in military vehicles: causes, consequences and cures*, pages 49–1—-49–8, 2002.

[170] Jan B. F. van Erp and Peter J. Werkhoven. Validation of Principles for Tactile Navigation Displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50:1687–1691, 2006.

[171] Alessandro Vinciarelli and Gelareh Mohammadi. A Survey of Personality Computing. *IEEE Transactions on Affective Computing*, 5(3):273–291, jul 2014.

[172] Patrik Vuilleumier and Yang-Ming Huang. Emotional Attention: Uncovering the Mechanisms of Affective Biases in Perception. *Current Directions in Psychological Science*, 18(3):148–152, jun 2009.

[173] Julia Wache, Ramanathan Subramanian, Mojtaba Khomami Abadi, Radu-Laurentiu Vieriu, Nicu Sebe, and Stefan Winkler. ASCERTAIN: Emotion and Personality Recognition using Commercial Sensors. *IEEE Transactions on Affective Computing*.

[174] Julia Wache, Ramanathan Subramanian, Mojtaba Khomami Abadi, Radu-Laurentiu Vieriu, Nicu Sebe, and Stefan Winkler. Implicit User-centric Personality Recognition Based on Physiological Responses to Emotional Videos. In *Proceedings of the 17th ACM on International Conference on Multimodal Interaction*, pages 239–246, 2015.

[175] Hee Lin Wang and Loong Fah Cheong. Affective understanding in film. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(6):689–704, 2006.

[176] Shang-fei Wang and Xu-fa Wang. Emotional Semantic Detection from Multimedia. In *Kansei Engineering and Soft Computing*, pages 126–146. IGI Global, 2011.

[177] Shangfei Wang and Qiang Ji. Video affective content analysis: a survey of state of the art methods. *IEEE Transactions on Affective Computing*, 3045(c):1–1, 2015.

[178] David Watson and Lee Anna Clark. The PANAS-X: Manual for the positive and negative affect schedule-expanded form. Technical report, The University of Iowa, 1999.

[179] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6):1063–70, jun 1988.

[180] Sidney Weinstein. Intensive and extensive aspects of tactile sensitivity as a function of body part, sex and laterality. In *the First Int'l symp. on the Skin Senses, 1968*, 1968.

[181] Christopher D. Wickens. Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2):159–177, 2002.

[182] Christopher D Wickens. Multiple Resources and Mental Workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3):449–455, jun 2008.

[183] Kathy A. Winter and Nicholas A. Kuiper. Individual differences in the experience of emotions. *Clinical psychology review*, 17(7):791–821, 1997.

[184] Charlotte V. O. Witvliet and Scott R. Vrana. Play it again Sam: Repeated exposure to emotionally evocative music polarises liking and smiling responses, and influences other affective reports, facial EMG, and heart rate. *Cognition & Emotion*, 21(1):3–25, 2007.

[185] W. Wundt. *Grundriss der Psychologie (Outlines of Psychology)*. Entgelmann, Leibzig, 1986.

[186] Xuehan Xiong and Fernando De la Torre. Supervised Descent Method and Its Applications to Face Alignment. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 532–539. IEEE, jun 2013.

[187] C. Y. Yaakub, N. Sulaiman, and C. W. Kim. A study on personality identification using game based theory. In *2nd International Conference on Computer Technology and Development*, pages 732–734. IEEE, nov 2010.

[188] Heng Yang, Changqing Zou, and Ioannis Patras. Face Sketch Landmarks Localization in the Wild. *Ieee Signal Processing Letters*, 21(11):1321–1325, 2014.

[189] Xiangkui Yao. Pedestrian Navigation Systems: a Showcase for Deep Personalization Theory. *Proceedings of the 9th international conference on Human computer interaction with mobile devices and services - MobileHCI '07*, pages 174–177.

[190] Nick Yee, Nicolas Ducheneaut, Les Nelson, and Peter Likarish. Introverted Elves & Conscientious Gnomes: The Expression of Personality in World of Warcraft. In *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, 2011.

[191] Lijun Yin, Xiaozhou Wei, Yi Sun, Jun Wang, and M.J. Rosato. A 3D Facial Expression Database For Facial Behavior Research. In *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, volume 2006, pages 211–216. IEEE, 2006.

[192] Gloria Zen, Bruno Lepri, Elisa Ricci, and Oswald Lanz. Space Speaks Towards Socially and Personality Aware Visual Surveillance. *ACM International workshop on Multimodal pervasive video analysis*, pages 37–42, 2010.

[193] Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang. A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Transactions on pattern analysis and machine intelligence*, 31(1):39–58, jan 2009.