UNIVERSITY OF THE WEST of SCOTLAND

UWS

Affective Computing in Controlled Exergame Environments

by

Larissa Müller

Thesis submitted in partial fulfilment of the requirements
of the University of the West of Scotland
for the award of Doctor of Philosophy

October, 2017

Table of Contents

Abstract

This work is located in the field of affective computing and relates to the design, construction and evaluation of an affective system. It gains insights into related affective computing research by providing an overview of emotion theories, current research approaches, emotion recognition and provocation technologies, as well as evaluation methodologies of affective systems.

A cycling exercise machine is enhanced and applied as a controller for a virtual racing exergame. The system setup provides a testbed for emotion recognition and emotion provocation techniques by means of a flexible and extensible architecture. Thereby a realistic testbed is applied for various affective real-world applications. Multimodal sensor data acquisition provides a more complete picture for the purpose of emotion recognition.

A virtual cycling game is designed to provoke specific emotions and steer participants in controlled emotional states. Single- and multi-emotion provoking game scenes are introduced for an enlarged study of diverse emotional provocations. The system setup is a controlled exergame environment to minimise irrelevant factors and create reproducible results.

A new affective game concept of a nonlinear game play is introduced to provide an enhanced gaming experience. The presented Emotional Journey is designed to allow a more individualised experience by means of a runtime adaptation to emotional reactions.

This work introduces novel event-based emotion analysis methods for more precise emotion recognition in an exergaming context. Two event-based emotion analysis methods are developed to recognise emotions from the facial expressions and physiological data of the participants while exercising. The combination of facial expression and EDA data analysis enhances the successful recognition rates. Both methods are enhanced by soft real-time abilities to enable a dynamic system response.

Two experiments are designed and conducted to validate, test, and evaluate the proposed methods and concepts. The first experiment with eleven participants is designed to showcase that the system is able to provoke user emotions. The second experiment with 25 participants evaluates new entertaining content, which includes multi-emotion provoking game scenes and physical stress. Furthermore, the presented Emotional Journey concept is evaluated and shows promising results for a predefined path through emotions. Finally, a database is created for the collected data of non-acted emotions.

Declaration

The research presented in this thesis was carried out by the undersigned. No part of the research has been submitted in support of an application for another degree or qualification at this or another university.

March 9, 2018	
Date	Signature

Acknowledgements

I am very grateful to everyone who helped me during my PhD studies. First of all, I would like to express my sincere gratitude to my supervising professor at the University of Applied Sciences in Hamburg, Prof. Dr Kai von Luck. His continuous encouragement and support throughout made this thesis possible. Furthermore, I would like to thank Dr Florian Vogt, whose support had a major impact on the success of this work. I am very grateful for his highly professional feedback, his guidance through the research process, and his patience. Moreover, I would like to thank the other, and not less important parts of my supervisory team, namely Prof. Dr Qi Wang and Prof. Dr Christos Grecos for their support, their valuable feedback and inspiring discussions. Thanks to Prof. Dr Qi Wang for becoming my new director of studies and to Prof. Dr Christos Grecos for his continuous encouragement from all over the world.

I would like to acknowledge Pro Exzellenzia for funding my doctoral studies and supporting my personal and professional development. Various workshops, coaching, and an excellent network have provided a solid basis for my future career. In addition, the graduate school at the HAW deserves a mention here, especially the support and advice of Prof. Dr Zita Schillmöller had a major impact on my studies, many thanks for that.

The EmotionBike platform was a joint project with Arne Bernin and Sobin Ghose. In addition, many students contributed to the success of the project. Therefore, I would like to thank the EmotionBike team for all the sleepless nights in which we worked together to keep deadlines, and for all the fun we had. Many thanks to Arne Bernin, who mainly developed the architectural and technical design, to Sobin Ghose for his contributions to the implementation and for his innovative ideas for the system architecture, to Sebastian Zagaria and Jörn Lambert for implementing the EmotionBike game, to Jonas Hornschuh for the technical implementation of steering capabilities, and to Thomas Berntin for designing and mounting the rotatable handlebar, to the Zentrale Laborwerkstatt at the HAW for their support, to Andreas Kamenz and Wojtek Gozdzielewski for implementing the physiological sensor data acquisition, to Florian Kletz and Jorin Kleimann for integrating the thermal camera, to Arne Bernin for benchmarking the facial expression algorithms, integrating video processing capabilities and the iMotions platform, to Erik Mathiessen and Kai Bielenberg for their technical support, to everybody who participated in one of the experiments or one of the many preliminary studies, and to Prof. Dr Kai von Luck, Dr Florian Vogt and Prof. Dr Thomas Netzel for supervising the project. I would also like to thank the faculty TI at the University of Applied Sciences for funding the whole project.

I would like to express my gratitude to Kai Rosseburg for providing photographs of the EmotionBike setup, to Dr Susanne Draheim for useful advice and to Ralf Jettke for valuable discussions. In addition I would like to thank Kai Rosseburg, Karolina Bernat, André Jeworutzki, Arne Bernin and Sobin Ghose for reviewing my work.

And finally, many thanks to Enno Putzar, my mom, my brother and all my friends as well as sports colleagues from the Taekwon-Do Black Belt Centre Hamburg for their patience and spiritual support during the PhD process.

List of Publications

Peer-reviewed Conference Papers:

1. Inproceedings: Larissa Müller, Sebastian Zagaria, Arne Bernin, Abbes Amira, Naeem Ramzan, Christos Grecos and Florian Vogt. EmotionBike: a study of provoking emotions in cycling exergames, International Conference on Entertainment Computing, 2015, 155-168;

2. Inproceedings: Larissa Müller, Arne Bernin, Sobin Ghose, Wojtek Gozdzielewski, Qi Wang, Christos Grecos, Kai von Luck and Florian Vogt. Physiological data analysis for an emotional provoking exergame, 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 2016, 1-8;

3. Inproceedings: Larissa Müller, Arne Bernin, Kai von Luck, Andreas Kamenz, Sobin Ghose, Qi Wang, Christos Grecos and Florian Vogt. Enhancing Exercise Experience with Individual Multi-Emotion Provoking Game Elements, 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017, 1-8;

4. Inproceedings: Larissa Müller, Arne Bernin, Kai von Luck, Andreas Kamenz, Sobin Ghose, Qi Wang, Christos Grecos and Florian Vogt. Emotional Journey for an Emotion

Provoking Cycling Exergame, 2017 IEEE 4th Intl. Conference on Soft Computing & Machine Intelligence (ISCMI 2017), 104-108;

List of Abbreviations

HCI	Human-Computer Interaction
HAW	University of Applied Sciences in Hamburg
EDA	Electrodermal Activity
ANS	Autonomous Nervous System
PAD	Pleasure Arousal Dominance
FACS	Facial Action Coding System
FER	Facial Expression Recognition
AU	Action Unit
SFB	Skin Blood Flow
PR	Pulse Rate
SC	Skin Conductance
GSR	Galvanic Skin Response
ECG	Electrocardiogram
BVP	Blood Volume Pulse
PPG	Photoplethysmogram
EEG	Electroencephalography
RPM	Revolution Per Minute

Table 0.1.: List of Abbreviations

List of Figures

List of Tables

1. Introduction

1.1. Background

Today, as predicted, ubiquitous computing (Weiser 1991) is on the rise by providing seamless interactions, digitalisation, and smart environments. This trend established a new connection between humans and computers, as it permits an enhanced relationship that is aligned to mirror relationships between humans. Computer science has a long tradition in redesigning this relationship, as building user models in dialogue systems is well known in artificial intelligence (Wahlster and Kobsa 1986).

The modelling of emotions in computer science enables new ways for user interface design by facilitating an additional level of interaction. For decades, cognitive scientists have been developing computer systems that model human emotions (Dorner and Hille 1995). Systems that are designed to have their own emotions have also been introduced, to provide a more natural counterpart in interactions (Becker and Wachsmuth 2006).

The research area that focuses on user emotions is called affective computing (Picard 1997a) and has yielded many applications to sense, model, express or provoke emotions during the last decade. The term affective computing was coined by Rosalind Picard, who describes her vision as follows: *"It is my hope that affective computers, as tools to help us, will not just be more intelligent machines, but will also be companions in our endeavors to better understand how we are made, and so enhance our own humanity."* (Picard 1997b)

Fundamental research, combined with technological progress, enables the establishment of new interfaces to evaluate the extended relationship between humans and computers. Research on emotional dialogues in human-computer interaction has gained more attention, inspired new ways of integrating emotions in interactive applications, and provided affective interactions (Peter and Beale 2008). This trend is accompanied by a great interest in the commercial sector of developing human-computer interaction technologies. Modern applications are constantly introduced, including various perception channels such as body sensors, gesture or speech recognition that can be applied to affective computing (Hook 2008). Many of them have already reached the consumer market, and various algorithms have been developed to analyse elaborate user actions. Therefore, modern technology provides enhanced methods for users to interact with computer systems compared to a classical desktop setting.

Cognitive sciences have shown that motivation accompanies emotions (Brandstätter et al. 2013). People react emotionally to events that affect them personally. These kinds of events often have a meaningful impact on a person's motivation, aims or needs. In addition, the wish for a positive emotional outcome or the avoidance of negative emotions can become a motivation in itself. A field that requires much intrinsic motivation is physical exercise. Computational assistance such as entertaining content has shown to provide motivation (Malaka 2014). Therefore, considering the relation between people's motivation and the corresponding emotions presents an interesting perspective as regards training or mobility applications. A physical exercise machine is also an interesting modern human-computer interface that is not a classical desktop setting.

This work combines affective computing technologies and physical exercise. The presented approach has great potential for the workplaces of the future, due to the rapid evolution from classical desktop setups to interactive and multimodal systems. A cycling exercise machine is applied as a controller for an emotion provoking game. Emotion recognition techniques are used to measure the provoked emotional responses. The concept of a dynamic game play is developed to steer the participants on a predefined emotional path by crafted emotion provoking game elements.

1.2. Motivation

Emotions play a very important role in everyday interactions between humans and in human-computer interactions (Peter and Beale 2008). Thus, the incorporation of affective computing leads to more harmonious human-computer interaction (Tao and Tan 2005). Many algorithms for emotion recognition have been developed, as various research fields benefit from the detection of user emotions. However, context is an important aspect for the reliable recognition of emotions (Westerink, Krans, and Ouwerkerk 2011). In this work, affective computing interactions are evaluated in a controlled setup. Exergames are applications that go beyond traditional computer use and enable the evaluation of new methods of interaction. Emotions can be provoked in a gaming scenario and combined with physical effort. Thereby, the entertaining content of exergames helps to improve fitness programmes by providing enhanced motivation, which supports a healthier lifestyle (Malaka 2014). In this work affective computing methods are applied to an exergame, to evaluate the potential of emotions to modern interfaces. As shown in trials in various scenarios and applications, adaptivity, such as the estimation of an adequate difficulty level that prevents frustration or boredom, is one important aspect of affective systems (Grafsgaard et al. 2013). Therefore, the presented emotion provoking game enables an adaptation to emotional reactions, by means of the designed and evaluated dynamic game play concept.

My personal motivation for the research in the field of affective computing is its interdisciplinarity. Among others, the field includes psychology, cognitive sciences, social sciences and computer sciences (Tao and Tan 2005), although this work is written under a computer science perspective. I am very interested in interdisciplinary research, and gained some experience during my master studies, shown in Müller et al. (2012a), Müller et al. (2012b), and Müller (2013). In a collaboration with a textile designer we developed interactive surfaces that were able to react to human emotions. During various exhibitions of our work we found that visitors were very interested in our "Emotional Dialogue" collection, hence I decided to continue researching emotions in HCI and build up on the great potential of affective computing.

1.3. Project Aims and Objectives

In this work a scenario is designed in which traditional methods cannot be applied, but enough fundamental research exists to provide an encouraging outcome. Concrete research questions will be clarified with the aid of an example to find generalisable and transferable results for the field of affective computing.

The applied approach is to combine techniques and algorithms that are able to sense user actions, as well as user activity, with current affective computing algorithms. This makes it possible to enhance the activity through the recognition and provocation of emotions. Many of the developed algorithms in affective computing are well tested and approved in a desktop scenario, but not in an interactive exercising context. Such a scenario requires movement and physical effort. It is interesting to analyse current research methods in this context, which go beyond the classical approaches of affective computing applications. The requirements for measurement systems differ on a technical and a conceptual level, and are thus intriguing to consider.

In this PhD thesis, current methods from affective computing will be applied in an exercise scenario. This requires a system that is able to provoke and measure user emotions during exercise. Moreover, the system should be able to adapt its provocation to the detected actions and emotions. Figure 1.1 illustrates the system components.

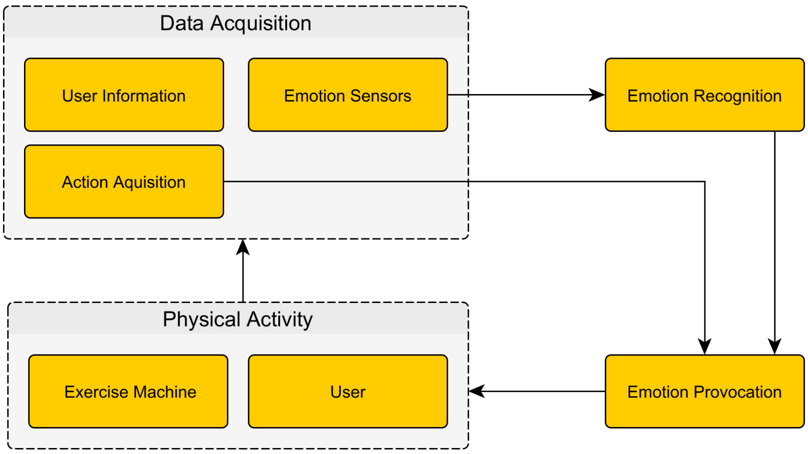


Figure 1.1.: Components of the presented Affective System

1.3.1. Research Questions

To specify the presented approach, the aim of this PhD thesis is to answer the research questions:

(1) Can a testbed be provided for various affective applications that provoke specific emotions in an exercise context?

(2) Can existing emotion recognition methods be applied to support reliable emotion recognition in an exergaming context?

(3) How can the concept of a dynamic system response to emotional user reactions in the testbed setup be applied and evaluated?

1.3.2. Research Objectives

The presented approach is also designed to provide a conclusion about the transferability and generalisation of the applied methods, and the results pertaining to the field of affective computing. This thesis will address the following objectives:

Objective 1: To gain insights into the field of affective computing and state of the art literature, which includes emotion theories and practical approaches of designing, building and evaluating affective systems.

Objective 2: To design and setup an affective exergame system with a loosely coupled architecture, which can be applied as a testbed for various applications and algorithms for emotion recognition and emotion provocation.

Objective 3: To provide an event-based analysis method with soft real-time abilities to recognise emotions through facial expressions and physiological data of participants while exercising.

Objective 4: To design a concept for a dynamic system response to emotional user reactions to steer the participants on a predefined emotional path by crafted emotion provoking game elements.

Objective 5: To design and conduct experiments to evaluate the emotional provocation of crafted game events, and the dynamic concept, by means of the developed analysis methods.

Objective 6: To provide a conclusion and discussion of the experimental results and their transferability and generalisation to the field of affective computing.

1.4. Technical Contributions

A cycling exercise machine is enhanced and applied as game controller for an immersive virtual cycling game. The system is designed to provide a testbed for emotion recognition sensors and algorithms as well as emotion provocation applications, to enable emotional provocation and the evaluation of participants' emotional reactions while exercising. A flexible and extensible architecture is applied, as multimodality is important for a reliable detection of emotions, and the open architecture enables an easy integration of different technologies and methods. Thereby an immersive setup is created that is able to provoke real and non-acted emotions, without distracting the player by difficulties in dealing with it.

An event-based method is developed to recognise emotions by analysing facial expressions and physiological data. Moreover, the event-based methods for emotion detection are enhanced by soft real-time abilities to enable a dynamic system response.

A virtual cycling game is designed to provoke specific emotions by tailored game elements. Thus, an experiment is designed and conducted to showcase the system's ability to provoke emotions. Moreover, the influences of personality traits on emotional responses is analysed.

The concept of adaptive game play is developed to steer the participants on a predefined emotional path by crafted emotion provoking game elements. An experiment is designed and conducted to evaluate the presented concept and crafted multi-emotion provoking game scenes. All experimental data are collected, thereby creating a database of non-acted emotions.

The results of the experiments are summarised, and the transferability or generalisation of the applied methods is discussed.

1.5. The Authors Role in the EmotionBike Project

The presented thesis is part of a research project at the University of Applied Sciences in Hamburg (HAW). The project idea has been developed by the supervising professor at the HAW, Prof. Dr Kai von Luck, the external supervisor Dr Florian Vogt, another PhD student Arne Bernin, and the author of this thesis. The project was funded by the faculty of engineering and computer science at the HAW. The project scope included budget for the technical setup and for two part-time jobs to maintain the system setup. The author of this thesis was funded by a scholarship; therefore another part-time employee was hired. Thereby the EmotionBike platform became a joint project between Arne Bernin, Sobin Ghose and the author of this thesis. Many students joined the team during the project duration. Some were involved in bachelor- and master projects, and others were paid to implement parts of the infrastructure or support the experiments. A detailed list of names can be found in the Acknowledgements of this work.

As a part of this thesis, an intense literature review is presented and research gaps in the field of affective computing are described. Based on these gaps, the thesis presents contributions to knowledge in Section 6.2; the corresponding system requirements are described in Chapter 3. The author's main challenge was to ensure that the system meet all requirements to enable it to recognise and provoke emotions in an exergaming setup, as well as to provide a dynamic system response to the emotional reactions. To evaluate the affective system the author designed and conducted two experiments, which are described in Chapter 4 and 5.

The design of the technical architecture of the EmotionBike system was mainly developed by Arne Bernin. A physical cycling exercise machine was enhanced by steering capabilities; the pedal resistance was made software controllable and transferred into a virtual cycling racing game. The author of this thesis defined the requirements described in Chapter 3 to provoke different specific emotions by software controllable game scenes. The author applied various tests to the technical setup and the physical exergame controller, because the physical setup should not influence the emotions of the participants during the experiments, for instance by insufficient controls.

One of the requirements was that all sensor data and the game events must be logged into the system to enable the analysis of emotional reactions. The author applied various tests to the emotion sensors and the data acquisition system. In addition to the need for reproducibility, the logging of all data was required to ensure the setup of a database with non-acted emotions. The author introduced a novel method for exergames to recognise emotions by facial expressions, as published in (Müller et al. 2015), and combined it with the analysis of physiological data as published in (Müller et al. 2016).

The requirements for the game design were defined by the author and are described in Chapter 3. The crafted game objects were chosen from freely available assets of the Unity Asset Store[1]. The technical implementation in Unity was carried out by a student team member, Sebastian Zagaria (first experiment). For the second experiment the requirements were enhanced by the author to ensure a dynamic system response and adaptation, including software controllable game events and a configurable difficulty. Jörn Lambert, a games master student at the HAW, was paid for the implementation of the game scenes of the second experiment. The author of this thesis tested the game scenes for both versions in preliminary studies with friends and students to improve the emotional provocations of the game scenes. Subsequent to those preliminary studies, the game scenes were refined in discussion with the respective developer.

The author introduced a novel concept of dynamic game play adaptation that ensures a predefined path of emotions based on using the designed game scenes, and published it in Müller et al. (2017a). Furthermore, multi-emotion provoking game scenes were designed and evaluated by the author and published in Müller et al. (2017b).

[1] https://www.assetstore.unity3d.com/en/

1.6. Thesis Structure

Chapter 2 summarises related affective computing literature. It provides an overview of emotion theories, and of methods to provoke and recognise emotions. Furthermore, it illustrates recent approaches of affective systems, and summarises experimental methodologies to evaluate such systems. At the end, identified research gaps are presented by differentiating between this thesis and other approaches.

Chapter 3 presents a new physical cycling game interface that provokes, records and analyses emotions. The loosely coupled system design supports an easy integration of new sensors for emotion recognition and thus provides a testbed for various affective technologies. Moreover, the design enables user studies, which are important for the evaluation of affective systems.

In **Chapter 4** an experiment is designed and a case study is described, which was conducted to demonstrate that tailored game events are able to provoke specific emotions. In addition, two novel event-based emotion analysis methods are presented. The first method is based on emotional reactions in facial expressions; the second enhances the facial expression analysis by physiological data.

In **Chapter 5** the analysis method is enhanced by real-time capabilities to enable a dynamic system response. It further illustrates new entertaining content, which contains multi-emotion provoking game elements. Moreover, it presents the results of a second case study that was conducted to evaluate the presented novel concept called Emotional Journey.

Chapter 6 summarises the thesis and discusses the transferability or generalisation of the applied methods. Furthermore, it highlights the main contributions to knowledge and future research perspectives.

Figure 1.2 illustrates the structure of this PhD thesis.

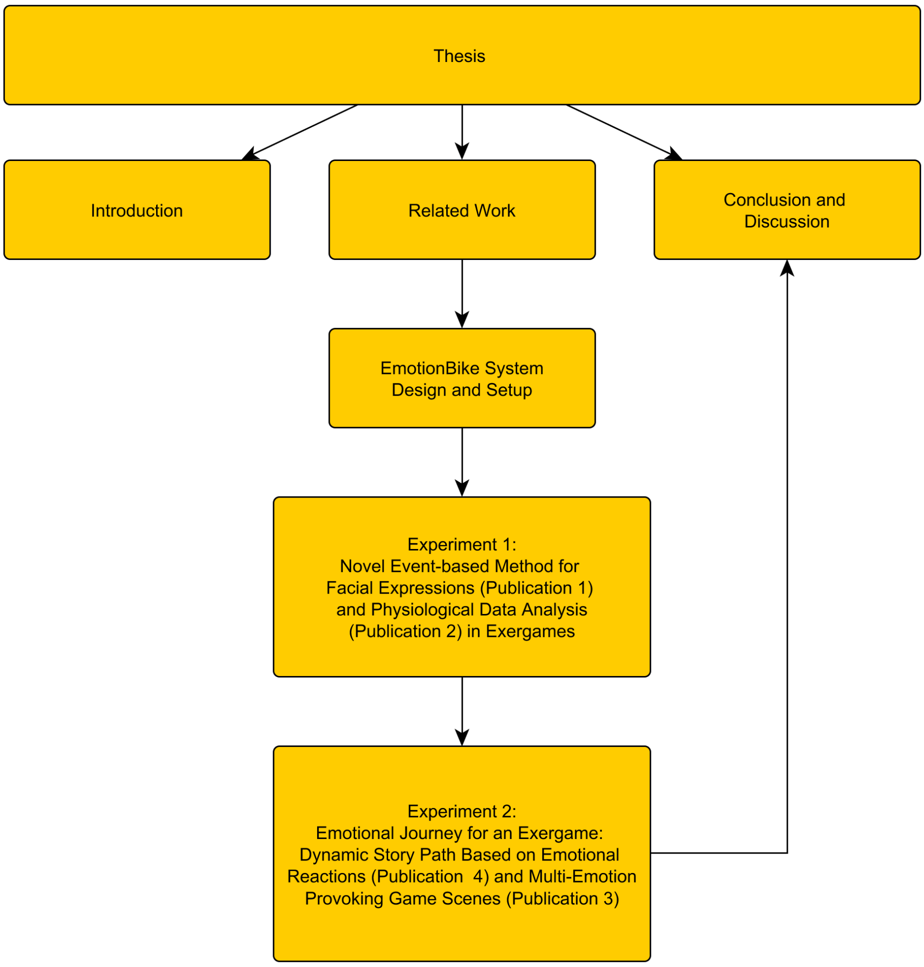


Figure 1.2.: Thesis Structure

2. Review of Related Literature

In this chapter an overview of recent methodologies in the field of affective computing is presented. The term "affective computing" was coined by Rosalind Picard (Picard 1997a) and according to Calvo and D'Mello (2010) her book was a landmark, which prompted a wave of interest in the research of improving human-computer interactions by means of emotions. Affective applications are designed to recognise and influence emotions, or to have their own emotions.

This PhD thesis is about designing, building and evaluating an affective system, as illustrated in Figure 2.1. The design of an affective system requires a suitable emotion theory (Section 2.1), as affective computing is an interdisciplinary field (Zeng et al. 2009). In addition, the emotional provocation (Section 2.2) needs to be defined to build an affective system. It is also important to have an overview of recent affective systems (Section 2.3), as the research field and its results are growing rapidly. New approaches have a significant influence on emotion recognition techniques, and reliable emotion recognition (Section 2.4) is essential for an affective system. Moreover, the thesis also presents an overview of state of the art affective computing experimental designs (Section 2.5), as the evaluation of affective systems is a challenging task.

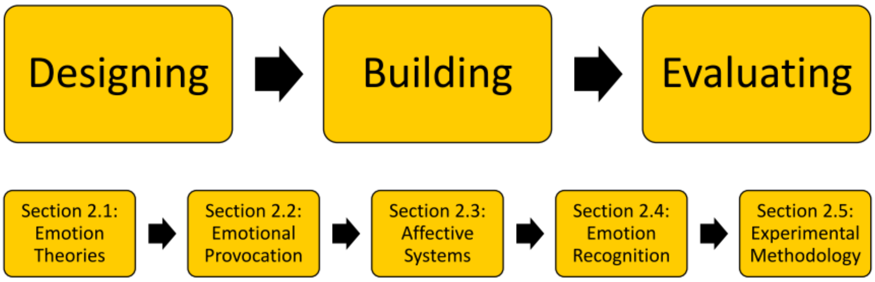


Figure 2.1.: Components of Related Research

2.1. Emotion Theories

An adequate emotion theory is essential for the design of affective systems (Calvo and D'Mello 2010). However, it is a very challenging task to find an appropriate theory in the highly interdisciplinary field of affective computing. Many terms are used synonymously, and sometimes even contradictory definitions can be found in related research. Moreover, often different emotion theories are suitable for an application. In the following, an overview of the relevant terms for this thesis and their definitions is provided, as well as selected emotion theories.

According to Scherer (2000) and Scherer (2005), the term emotion is a hypothetical construct that is hard to define. He also claims that it has often been used incorrectly, for instance by James (1884) in his work "What is an emotion", as that relates to feelings rather than to emotions. A review of various emotion definitions was presented by Kleinginna Jr and Kleinginna (1981). *"Emotions, feelings, moods, attitudes, affective styles, and temperament"* were described as relevant affective phenomena in Calvo and D'Mello (2010), who also provide a comprehensive overview of the modelling of affects in the research field of affective computing. In the following relevant terms for this thesis are presented: **Feeling**, **Affect**, **Emotion**, **Expression**, **Appraisal** and **Mood**.

Feeling: Scherer (2000) described the term Feeling as a subjective experience of emotional arousal.

Affect: According to Scherer (2000) the words Emotion and Affect are used synonymously, but by some researchers they are restricted to the valence aspect of feelings (e.g. positive vs. negative). In this work both terms are used synonymously, but focus mainly on the term Emotion.

Emotion: Calvo and D'Mello (2010) describe Emotions as very complex and fuzzy, as well as indeterminate and elusive constructs. According to Hakim, Marsland, and Guesgen (2013), *Emotions* are dynamic and vary continuously in intensity, duration and persistence with time.

Expression: According to Ekman and Friesen (1976) and Scherer (2000), Expressions are muscular actions in for instance someone's face, vocal cords or hands.

Appraisal: Scherer (2000) describes Appraisal as an evaluation of the significance of an event.

Mood: According to Scherer (2000) Mood lasts longer than Emotions. In addition, Mood is not necessarily related to a specific event. According to Becker (2003) a person's Emotion is strongly related to a person's Mood, thus a person who is in a positive Mood will not react as aggressively as if he or she was previously in a negative Mood.

Many related works focus on the term Emotions (Calvo and D'Mello 2010), others apply the term Affect instead of emotion (Zeng et al. 2009). This PhD thesis focus on the term Emotion. The practised terminology in related research is often intertwined with the emotion theories beyond research. All emotion theories are very controversial in psychology, and according to Hakim, Marsland, and Guesgen (2013) there is still no objective definition of emotions in psychology. A classification of recent methodologies in emotional modelling is thus a challenging task. The presented thesis is written by a computer scientist and thus no judgement will be made. However, an overview about relevant models is presented in the following, due to its importance for designing affective systems.

Two computational views on emotional modelling are outlined: **Discrete Modelling of Emotions** and **Dimensional** or **Continuous Modelling of Emotions**. The underlying concepts of these two ways of modelling have been adapted and applied in other emotion theories as well. Psychological categories have for instance been described by Calvo and D'Mello (2010), who presented six theoretical perspectives on emotions: **Emotions as Expressions**, **Emotions as Embodiments**, **Cognitive Approaches to Emotions**, **Emotions as Social Constructs**, **Neuroscience**, **Core Affect and Psychological Construction of Emotion**.

Emotions as Expressions includes the theories of Darwin (1872), that emotions in the human's face are universal and inherent. Furthermore, it contains the discrete emotional modelling of basic emotions proposed by Ekman and Friesen (1976). Thus, in this thesis it will be called **Discrete Modelling of Emotions** and be thereby described as a computational view. Calvo and D'Mello (2010)'s description of **Core Affect and Psychological Construction of Emotion** includes related research, which is described as **Dimensional Modelling of Emotions** in this work, and thus is also described as a computational view. Furthermore, the theories categorised as **Emotions as Embodiments** are called **Body Response Theories**, as in this work they also describe theories of body changes influencing emotions. Moreover, the categories **Emotions as Social Constructs** and **Cognitive Approaches to Emotions** will be summarised due to their relevance for this work. Details of **Neuroscience** aspects are not provided, as they are not the focus of this PhD thesis.

In psychology, the theories of emotions are very controversial. Calvo and D'Mello (2010) complain about the fact that a best practice of modelling has not yet been developed. The approach of modelling an emotion as a point in a valence arousal space is as acceptable as a categorisation to predefined labels. Picard (1997a) herself advocates pragmatism in setting up affective systems.

2.1.1. Discrete Modelling of Emotions

In his work "The expression of the emotions in man and animal" Darwin (1872) hypothesised that the expressions of emotions in the human face are universal and inherent. Ekman (2004) stated that he started his work with the idea of disproving this theory, because he believed that emotions are culturally appropriated. However, it was not possible for him to disprove it. On the contrary, as a result of his research with aboriginal people in Papua New Guinea, Ekman (2004) found that six basic emotions are universal. He showed pictures of emotional facial expression of Caucasian people to Papua New Guineans and vice versa. He claimed that six emotional expressions are universal: happiness, sadness, surprise, disgust, anger and fear, as illustrated in Figure 2.2. Therefore, he defined a theory of six **Basic Emotions** expressed by facial expressions. Furthermore, he asserts that there are unique expressions for the different emotions, which he used to explain that the facial muscle movement by the expression of emotions is similar for all humans. Many works are based on the six basic emotions: anger, disgust, fear, joy, sadness and surprise, although Ekman (2016) raised the number to seven by adding contempt.

Figure 2.2.: Basic Emotions

Ekman and Friesen (1976) proposed the Facial Action Coding System (FACS), a system to measure facial movements, called Action Units (AUs). Trained FACS coders are able to categorise basic emotions on the basis of AUs expressed in a human's face. In the last decades computer vision methods have become more sophisticated, and applied FACS to support a non-intrusive way of emotion detection. Nowadays there are various systems to measure facial expressions, which will be further described in Section 2.4.1.

An advantage of the theory of basic emotions is that a discrete labelling scheme matches the experiences of people, as they often fail to describe every range of emotion (Zeng et al. 2009). It is a frequently applied emotion theory in affective computing, as the recognition of basic emotions is very common in related literature (Stratou et al. 2013; Hakim, Marsland, and Guesgen 2013), although, as mentioned, all theories are controversial. One major criticism of the basic emotion theory is that such a discrete categorical approach may not reflect the complexity of affective states such as "thinking, embarrassment or depression" (Gunes et al. 2011a).

To summarise, Ekman and Friesen (1976)'s theory of universal emotions expressed by facial expressions supports a non-intrusive way of emotion detection, and there are well-studied systems to measure facial expressions. Therefore, they are part of this PhD research. More details are presented in Section 2.4.1.

2.1.2. Dimensional Modelling of Emotions

According to Becker-Asano (2008) the first ideas of a dimensional view on emotions were proposed by Wundt (1863) and includes three axes, ranging from pleasure to displeasure, excitement to inhibition, and tension to relaxation. A widely used dimensional model, the "Circumplex Model of Affect" was proposed by Russel (1980) and describes a 2D way of modelling emotions, as presented in Figure 2.3. The arousal axis ranges from inactive to active and measures how energised or soporific one feels. The valence axis ranges from unpleasant (e.g sad, stressed) to pleasant (e.g. happy). Garbas et al. (2013) describe this model as an attempt to map emotion labels in a two-dimensional continuous emotion space, and mention that it was enhanced by Scherer (2005). Garbas et al. (2013) utilised a valence-based approach, using the classifications by Russel (1980) and Scherer (2005) to divide training data of available databases into positive and negative valence. They also cluster faces without emotional responses present as neutral, and label the emotions: happy, angry, disgusted, sad, frightened, surprised. However, they excluded surprise, based on the findings of Russel (1980) that it is neither positive nor negative. The work of Garbas et al. (2013) is located like many other affective computing systems in the area of market research. Market research setups often include a static desktop scenario and thereby they do not require a natural interaction, as it was required in this work.

Related works use the term activation-evaluation space (Hakim, Marsland, and Guesgen 2013) or valence-activation space, which is similar to valence arousal (Gunes et al. 2011b). Figure 2.3 illustrates the two axes of valence and arousal, which are the most commonly applied dimensions.

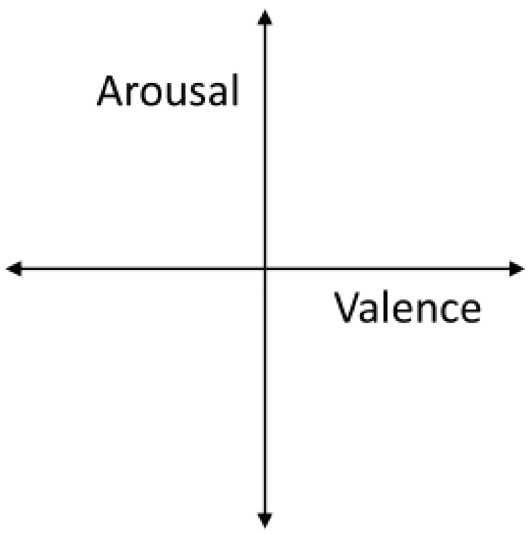


Figure 2.3.: Dimensional Modelling of Emotions

Plutchik presented a wheel of emotions that consists of the eight fundamental emotions: joy, surprise, sadness, disgust, anger and anticipation (Plutchik 2001). It also includes more advanced affect words and is used to illustrate different emotions in a nuanced way. It includes the theory that primary emotions can be expressed at different intensities and can mix with one another and thereby form different emotions. Another two-dimensional model has been introduced by Thayer (1990). One dimension displays energy and the other describes tension.

A third dimension of dominance can be added to the two-axes model, which ranges from feeling helpless to feeling empowered (Koelstra et al. 2012). The combination of the axis pleasure arousal and dominance is often called **"Pleasure Arousal Dominance (PAD) space"** and was introduced by Russell and Mehrabian (1977). Various related literature applies the PAD space to model emotions (Becker-Asano 2008; Soleimani and Kobti 2013).

A major criticism is that the dimensional approach is not intuitive for raters, as it allows to label a range of emotions (Zeng et al. 2009).

2.1.3. Body Response Theories

According to Calvo and D'Mello (2010) it was James (1884), who linked the theory of expression with physiology. The **James-Lange Theory** describes emotions as changes in the sympathetic nervous system, which is a part of the autonomic nervous system (ANS). According to Brandstätter et al. (2013), Lange (1887, Danish org. 1885) published the same ideas almost at the same time as James (1884). Thus, the theory includes the name of the two authors William James and Carl Lange. The theory states that physiological arousal instigates the experience of an emotion, thus the physiological reaction is prior to the brain's perception. It can be described as a person feeling sad because he or she is crying. This is a critical point of this theory since it is also possible that a person cries because he or she feels happy, therefore more information is required.

Many researchers in the field of affective computing apply physiological sensors, as physiological changes have been shown to be related to emotions (Koelstra et al. 2012; Kindness, Mellish, and Masthoff 2013a). Moreover, Picard and Daily (2005) claim that physiological measurements allow to observe affective states, even without relying on personal judgements of emotional states.

According to Brandstätter et al. (2013) a further theory was developed in the 1970's: the **neo-Jamesian theory** investigated the influence of changes on mimics and gestures on emotions. For instance, Strack, Martin, and Stepper (1988) showed that mimic changes have the ability to influence emotions. In their experiment they asked participants to put a pencil between nose and lips, which affects the same muscles as a natural smile, and presented them with emotional stimuli (they used funny cartoons). Their results showed that the affective experience was influenced; the participants with the pencil rated the cartoons as funnier than the control group.

Furthermore, there is related research applying physiological measurements, which is focused on **stress** because it is asserted to negatively influence health (Kudielka et al. 2007; Organization 2001; Giraud et al. 2013b), well-being and social interactions (Sharma et al. 2013). McCrae (1990) defined stress as *"[...] a burden (or strain) placed upon individuals by external events or conditions that tax their psychological capacities to adapt"*. The American Psychological Association (last access: 2016-09-05) classified stress as: acute, episodic and chronic. They define acute stress as short-term reactions to events, and mention that it might be perceived as thrilling and exciting in small doses. Episodic stress is described as episodes of acute stress. If stress is perceived over a long time it is classified as chronic stress.

Electrodermal Activity (EDA) increases with stress and anxiety but varies linearly with overall stress levels, while heart rate increases with arousal (Kindness, Mellish, and Masthoff 2013a). Moreover, it has been shown that stress can increase blood pressure (Koelstra et al. 2012). A low heart rate can be More details can be found in Section 2.4.2.

2.1.4. Social Constructs in Emotion Theory

According to Calvo and D'Mello (2010) it was Averill (1980) who claimed that emotions cannot be explained by physiological or cognitive terms only. He postulated that emotions are also related to social constructs. Opponents of the emotions as expression theory often utilise this argument to question Ekman and Friesen (1976)'s theory of basic emotions described in Section 2.1.1.

Sauter et al. (2010) investigated cultural differences in the intensity of emotional expressions, and Scherer (2000) mentions cultural differences in event appraisal. Additional arguments to strengthen this theory have been published in linguistic sciences, as there is no adequate translation for each emotion in different languages (Russell and Sato 1995). Intercultural differences have even been reported in research of social robots (Becker-Asano and Ishiguro 2011). Social influences are well known to effect emotional reactions. For instance, it has been shown that the perception of being judged instigates a social-evaluative threat, which induces stress (Dickerson and Kemeny 2004). The presented PhD thesis capitalises on these effects to intensify the emotional provocation. Very challenging tasks are tailored to induce ambition, which is often based on socially related emotions like for instance proud or shame.

2.1.5. Cognitive Approaches

According to Calvo and D'Mello (2010), Arnold (1960) is the pioneer of cognitive approaches to emotions. Reisenzein (2006) also describes her as the pioneer of modern cognitive emotion theory in psychology. Her theory is based on the idea that an event must be appraised or evaluated in order for a person to experience an emotion. The events are examined for instance regarding their ability to cope, their novelty or urgency, as well as their consistency to the goals (Calvo and D'Mello 2010). Summarising the personal appraisal of an event causes an emotional reaction.

Ortony, Clore, and Collins (1988) state that cognitive scientists assume that emotions arise as an effect to a situation in the way it is construed by a person. As an example they mention the results of a sports competition, where one team will be happy after the game but the other team will have negative feelings. They proposed the **OCC Model**, which is commonly used in related research to model emotions as cognitive appraisal (Calvo and D'Mello 2010). It is highly applicable for emotion modelling of virtual agents (Steunebrink, Dastani, and Meyer 2009). It structures emotions in three types: events based, agent based or object based. An event can be perceived as pleasant or unpleasant and can have consequences that can affect the person himself or others, such as for instance the emotions of joy and pity. Actions of agents can be construed as approving or disapproving and can lead to pride in case of a self-agent and to admiration if it is focused on another agent. Objects can be liked or disliked.

Another cognitive approach is **The Component Process Model of Emotion** by Scherer (2005), which describes emotions as changes in five subsystems of an organism in response to the personal evaluation of an event. Assuming that the stimulus event is relevant for a person, the emotion components include: appraisal, bodily symptoms, action tendencies, motor expressions (facial, vocal) and emotional experience. According to Zeng et al. (2009), Scherer's appraisal-based approach is the most influential in modern psychology. Becker-Asano (2008) developed the computational model of emotions (WASABI) for a virtual avatar, which allows to model an avatars mood and emotions and uses the PAD space. Furthermore Becker-Asano and Wachsmuth (2010) distinct between primary and secondary emotions. Primary emotions are expressed as immediate responses to a stimulus. Secondary emotions arise from a higher cognitive process which is learned in the social environment (Becker-Asano and Wachsmuth 2010; Becker and Wachsmuth 2006).

2.1.6. Summary of Emotion Theories

The terms **Affect** and **Emotion** are used synonymously in related research. The main categories of a computational representation are the **Discrete Modelling** of emotions and the **Dimensional Modelling** of emotions. The description of emotions in affective computing research as a point in a dimensional space is as acceptable as it is with predefined labels.

The main psychological theories are illustrated in Table 2.1. The theory of **Basic Emotions** describes emotions as universal and inherent, while the **Body Response** theory defines emotions as being related to physiological reactions. The theory of emotions as **Social Constructs** is based on the idea that emotions appears in interactions with others. The **Cognitive Approaches** additionally include personal experiences. All of the theories are highly controversial in related research, however in this work no judgement is provided, as all theories have a justification. The provocation and recognition of emotions is independent from the underlying emotion theory, although the **Body Response** theory is often applied

	Theory or Representation	Emotion Processing	Pioneer	Emotion Trigger
Discrete	representation		Darwin (1872)	
Dimensional / Continuous	representation		Wundt (1863)	
Basic Emotions	psychological theory	universal and inherent/innate	Ekman and Friesen (1976)	
Body Response	psychological theory	evolutionary	James (1884) and Lange (1887, Danish org. 1885)	closely related to physiological reaction
Social Constructs	psychological theory	interaction with others	Averill (1980)	social context dependent
Cognitive	psychological theory	based on personal experience	Arnold (1960)	situational self-awareness

Table 2.1.: Psychological Theories and Representation

for physiological data analysis and the **Basic Emotions** theory for facial expression recognition.

2.2. Emotional Provocation

The provocation of emotions is an important part of some affective systems, thus an overview of related research is provided. Different terms for **Emotional Provocation** exist. Wang and Marsella (2006) for instance call it **Emotion Evoking**, however in this PhD thesis the term **Emotional Provocation** is applied.

Various stimuli have been applied to provoke emotions, and an overview is presented in this section. Affective systems with a focus on the detection and provocation of stress often apply the term stressor instead of stimulus and will also be presented.

2.2.1. Stimuli

A stimulus that has been used to provoke emotions in related literature is **haptic**. Tsalamlal et al. (2013) for instance, propose a haptic device for the tactile stimulation with a mobile air jet. The author developed new emotional interfaces in previous interdisciplinary works with textile designers. These interfaces mapped sensed emotions based on facial expression recognition to an abstract device. For instance, the emotion sensitive active surfaces called "Lomelia", shown in Figure 2.4, expressed emotions by changing its physical attribute and was designed to involve the viewer in an emotional dialogue. Further details can be found in Müller et al. (2012a), Müller et al. (2012b), and Müller (2013).

Figure 2.4.: Emotion Sensitive Active Surface: Lomelia

Other commonly applied stimuli are **video or movie clips** (Munia et al. 2012; Cheng and Liu 2013; Abadi et al. 2013). It has been shown that fast moving scenes or objects provoke excitement (Koelstra et al. 2012). Jang et al. (2011), for instance presented emotional video clips from films or sport shows with a length of 2-4 minutes. Sharma et al. (2013) used film clips to induce stress and Sharma and Gedeon (2013) provided stressful film clips to provoke fear and tension combined with non-stressful films of mediation.

Music videos are used as stimuli to provoke emotions (Koelstra et al. 2012). It has been shown that **audio** or acoustic signals are good for provoking emotions (Chittaro and Sioni 2013). In addition, **instruments** have been utilised in affective computing research (Glowinski et al. 2013), and Alaoui-Ismaïli et al. (1997) utilised **odorants** as stimuli for basic emotions.

Mixed media, for instance a power point presentation have been utilised as stimuli by Munia et al. (2012), who additionally applied a **question answer method** in which participants were asked to memorise emotional situations. Moreover **virtual reality** (Krieger, Lallart, and Jouvent 2013) has been used as stimuli. Nasoz, Lisetti, and Vasilakos (2010) for instance developed a virtual reality environment designed as a car interface.

Other commonly applied stimuli in related research are **images** (Wang and Marsella 2006). The selection is often based on the International Affective Picture System (IAPS) (Garbas et al. 2013), which was proposed by Lang, Bradley, and Cuthbert (2005).

It has been shown that **video games** are an excellent way to provoke emotions, as they allow to trigger task-related emotions such as frustration (Wang and Marsella 2006).

2.2.2. Stressor

Related research with a focus on stress recognition uses the term **stressor** instead of **stimulus**. According to Kindness (2013) and Lazarus (1999), any stimulus that results in a stress response is known as a stressor. Hong, Ramos, and Dey (2012) claimed **mental arithmetics**, **loud sounds** and **cold water pressor** are common stressors. They analysed physiological reactions to these stressors during physical activity. They did not include facial expressions recognition although they might also be influenced.

Akhanda, Islam, and Rahman (2014) utilised **mathematical problem-solving tasks** and **memory tasks** as stressors. Furthermore Singh, Singla, and Jha (2009) conducted a case study that included **mental tasks** for thinking, long term memory activation, moving and emotions. As emotional stressor they applied an aptitude problem.

A further commonly utilised stressor is **public speaking tasks** (Feldman et al. 1999; Pfister and Robinson 2011). Those tasks are often based on the Trier Social Stress Test, which was introduced by Kirschbaum, Pirke, and Hellhammer (1993). It consists of a simulated job interview with the assignment to prepare a speech and present it in front of a jury of experts. Subsequently the jury provides positive and negative feedback. It has been shown that the test is successful in stress provocation by Kudielka et al. (2007). The test is often adapted (Giraud et al. 2013b), for instance by Chellali and Hennig (2013), who utilised a high stake public speaking task, a thesis defence, for their study. In psychology Dickerson and Kemeny (2004) showed that social-evaluative threats are perceived as stressful, which is related to the expectation that a bad performance will be negatively judged by others.

Video games have been applied as a stressor. It has been shown that stress can be induced by negative feedback in gaming scenarios (Wallbott and Scherer 1991), but the results are controversial (Friedrichs et al. 2015). In the presented PhD thesis game scenes are particularly tailored to provoke stress.

2.2.3. A Controlled Exergame Environment to Provoke Emotions

It has been shown that games offer a sophisticated and systematic way of provoking emotions through game events, due to a fully controllable environment (Wang and Marsella 2006). In this PhD thesis the advantages of a controllable game environment are combined with a controllable physical cycling exercise machine. Thus, physical effort can be applied as a supplementary opportunity of provoking emotional responses in the presented exergame.

The term **exergame** is a portmanteau of "exercise" and "gaming" (Hoda, Alattas, and El Saddik 2013). Games that focus on exercises are also called **exertion games** in related research (Chatham and Mueller 2013; Walmink, Wilde, and Mueller 2014). However, in this PhD thesis the term **exergame** is applied.

Much research in the field of exergames is motivated by the idea to fight the worldwide obesity problem, which can lead to health problems such as diabetes, cardiovascular disease or high blood pressure. Furthermore, it has been shown that exergames have the potential to encourage a healthy and active lifestyle (Hoda, Alattas, and El Saddik 2013; Malaka 2014). A balanced workout programme and adequate nutrition are very important today, due to the increasingly sedentary lifestyle (Hamilton et al. 2012). Accordingly, physical activity improves physical health (Lathia et al. 2016).

Motivation is crucial for exercise (Süssenbach et al. 2014) and entertainment can be highly motivating. Warburton et al. (2007) found evidence that combining video games and indoor cycling leads to a greater positive impact on health than other types of cycling. Accordingly, a combination of indoor cycling and video games is able to increase physical fitness (Hoda, Alattas, and El Saddik 2013; Warburton et al. 2007). Moreover, Hoda, Alattas, and El Saddik (2013) found that the combination of a game with a physical bike has a significant effect on speed, average rotation per minute and heart rate by motivation, resulting in an improved exercise. Furthermore, they described a better user experience as a key success factor of exergames and found that exergames are more entertaining and more encouraging to people with a basic activity level. They used a casual exercise machine which does not support any steering capabilities. The game experience could be improved by more degrees of freedom in interacting with the game environment. Furthermore they focus on heart rate data and did not analyse the emotional reactions in the participants faces.

Biddle, Fox, and Boutcher (2000) linked physical activity to emotion and mood. Lathia et al. (2016) also found that people are happier while exercising physically. Thus, physical activity has a positive effect on a person's well-being. As a matter of fact, Süssenbach et al. (2014) claimed sports to be crucial for health and well-being, and Lathia et al. (2016) found that people are happier if they are more physically active. However, another interpretation might also be true; happier people are more physically active.

In this work, new entertaining content is illustrated to support an exciting experience through gaming and exercising. The presented controlled exergame environment enables the provocation of stress or specific emotions through game elements that are based on a combination of mental and physical stimulation.

2.2.4. Summary of Emotional Provocation

Various stimuli have been utilised to provoke emotions, which are natural and spontaneous, as they are different to acted emotions. In stress related research, the term stressor is often used. Table 2.2 illustrates an overview of emotional provocation in related research, and shows that many stimuli or stressors are based on a static perception. Others are correlated with task performance of a choreographic process. There are only a few related works that combine perception and performance aspects in the emotional provocation and even less supports interactive and adaptive feedback or enhance their affective system with physical activity.

	Stimuli/ Stressor	**Physical Activity**
media perception	haptic	
	odorants	
	audio	
	video clips	
	music videos	
	images	
	mixed media	
	Virtual Reality	
task performance	mathematical problem solving	
	memory tasks	
	mental tasks	
	public speaking tasks	
	question answer method	
	play instrument	X
	sport	X
perception and performance	video games	
	exergames	X

Table 2.2.: Emotion Provocation Overview

2.3. Affective System Approaches

Affective systems have been shown to enhance the quality of human-computer interaction by making it more usable, enjoyable and effective (Calvo and D'Mello 2010). Affect-sensitive tutoring systems for instance can be enhanced by an adequate reaction to students' frustration to increase motivation and improve learning gains (Sidney et al. 2005; Schaaff and Adam 2013; Nasoz and Bayburt 2009), as it has been shown that affective states influence learning outcomes (Grafsgaard et al. 2013).

Affective learning systems often focus on emotions like confusion, frustration, boredom (Calvo and D'Mello 2010) or even anxiety (Nasoz and Bayburt 2009). Grafsgaard et al. (2013) for instance focused on frustration detection by analysing relationships in automatically detected Action Units (AUs), affective outcomes, and learning gains in a corpus of computer-mediated human-human tutoring. Sidney et al. (2005) developed a tutoring system (called AutoTutor) that classifies emotions by facial expressions, body movement and speech analysis. Other related research of affective learning systems has focused on providing appropriate feedback to increase motivation by considering the user's personality. Dennis, Masthoff, and Mellish (2013) developed a system providing feedback about learning performance combined with emotional support.

Affect detection has been applied in various areas. It has been claimed to support **hyper-narrative movies** (Abadi et al. 2013) and has been utilised for **mimicry recognition** (Bilakhia, Petridis, and Pantic 2013). It has also been used in the fields of medical and fitness applications, for instance to assist psychologists in the **automatic recognition of depression**. Alghowinem et al. (2013) found by analysing head pose and movement that depressive persons did not change their head position as often as a control group. Moreover, they move their head more slowly and spend more time looking down or to the right. The authors conclude that depressed persons might avoid possible eye contact, or that they are tired. Stratou et al. (2013) found gender relevant for an automatic assessment of depression and post-traumatic stress disorder (PTSD). In their study women showcased more emotional variability than men in distressed and non-distressed conditions.

Related affective computing research has shown that the perception of **virtual avatars** by humans changes if avatars simulate human-like behaviour (Becker-Asano 2008). The agents are assumed to be more trustworthy and believable if they, for instance, apologise (Becker et al. 2005). Embodied conversational agents have been developed that are able to display empathy or engagement, motivate people and give confidence (Gunes et al. 2011b). Other works focus on a more natural appearance of virtual avatars. Ochs, Prepin, and Pelachaud (2013) facilitated a more natural smile, and Amini and Lisetti (2013) presented an open source software to generate realistic facial expressions on 3D virtual characters. Moreover, Liebold and Ohler (2013) found that multimodality enhances the user perception. Their results showed that virtual avatars that use multimodal expressions such as facial expressions and prosody have higher recognition rates.

Affective computing research has already advanced from virtual avatars to **physical androids** (Becker-Asano 2011). Human replicas that appear almost, but not exactly, like real human beings often provoke feelings of eeriness or revulsion, described as Uncanny Valley (Mori, MacDorman, and Kageki 2012, Japanese orig. 1970). Thus, related research forces the development of human-like facial expressions (Becker-Asano and Ishiguro 2011) or laughter (Becker-Asano et al. 2009), to overcome these effects (Becker-Asano et al. 2010).

The research of human-like robots has reached a stage where the robot itself is not even recognised as a robot in certain circumstances. In a case study Pütten et al. (2011) showed that persons who had no direct eye contact with the robot often mistook the android for a human. Mazzei et al. (2012) in fact applied an android in behavioural studies on children with autism.

Dialogue systems that support a natural verbal interaction are essential for human-robot interaction. The robot head "Flobi" (Lütkebohle et al. 2010) for instance has been used to play the game "pairs" with humans by using a dynamic dialogue system (Kipp and Kummert 2014). Multimodality is crucial for such systems. Kennington, Kousidis, and Schlangen (2014a), Kennington, Kousidis, and Schlangen (2014b), and Baumann and Schlangen (2012) for instance presented an architecture that makes it easy to plug in different sensors. They use the Google web speech API, a Leap motion sensor[1], a Microsoft Kinect[2] camera and an eye-tracker. They applied their dialogue system for driving scenarios and tracked driver attention (Kousidis et al. 2014).

Related research in the area of **fitness applications** was introduced by Süssenbach et al. (2014). They developed an indoor cycling companion to motivate persons during their exercises by utilising a NAO[3] robot.

Another field of related research is **biofeedback systems**. These systems provide feedback about the measured physiological data to the user, for instance to cause a calming effect. However, the results are highly controversial, since for instance Raaijmakers et al. (2013) found no significant effects by evaluating skin conductance and heart rate variability.

[1] https://www.leapmotion.com/
[2] https://developer.microsoft.com/en-us/windows/kinect
[3] https://www.ald.softbankrobotics.com/en

Walmink, Wilde, and Mueller (2014) in turn developed a bicycle helmet that displays heart rate data and found that it results in a social interplay and facilitates engagement with the exertion activity.

One industry that benefits significantly from affective computing research is the entertainment sector. Thus, many commercial games (Christy and Kuncheva 2014) have been developed during the last decade in the field of **affective games**. Affective games are described as a form of gameplay where the current emotional state of a player is used to alter the game mechanics (Gilleade, Dix, and Allanson 2005). Affective games often use biofeedback in their design, but according to Gilleade, Dix, and Allanson (2005) a game can also be affective without offering biofeedback. Wang and Marsella (2006) for instance introduced the Emotion Evoking Game (EVG), a dungeon role playing game that is designed to provoke emotions. Their aim was to support the development and analysis of new techniques for recognising emotions and generating human-like facial expressions for virtual avatars. Their ideas are based on the Geneva Appraisal Manipulation Environment (GAME) project (Kaiser and Wehrle 1996), which is a tool to generate experimental Pacman-type computer games. In their case study Wang and Marsella (2006) offer different **emotional events** to the participants. Their sequence started with inducing boredom, followed by surprise, joy, anger and disappointment. In addition they claim that social emotions like guilt can be triggered. They did not provide a dynamic game play which would provide an improved game experience.

Vachiratamporn et al. (2014) applied a survival **horror game** as experimental environment for an affective game. They analysed heart rate and electroencephalography (EEG) data and changed the timing of horror events based on the player's affective states, but found no significant differences between in the EEG data for the affective and non-affective version of the game. Their approach is based on a static desktop scenario of a person sitting in front of a computer, therefore their analysis methods cannot easily be applied in exergaming scenarios.

Negini, Mandryk, and Stanley (2014) adapt the **game difficulty** in a first-person shooter regarding Galvanic Skin Response (GSR) data and found that affectively adapted games enhances participants' arousal. They also evaluated different strategies of game manipulation and found that supportive strategies, e.g. giving the player tools needed to overcome challenges, are more effective than changing the strength of enemies. Parnandi, Son, and Gutierrez-Osuna (2013) presented a racing game that adapts game difficulty automatically based on physiological data analysis and used speed, road visibility in weather conditions and steering jitter. They mention that an adaptive affective game should be intuitive, engaging, and amendable to adaptation, and describe the advantages of a game of the car-racing type as intuitive, easy to learn, highly dynamic and enabling multiple forms of adaptation. The EmotionBike game is designed as a bicycle fun-racer, which is similar to **car-racing games**. However the introduced dynamic system control goes further than changing game difficulty. Instead the story path is adapted to the measured affective state.

2.3.1. Summary of Affective System Approaches

This section presented affective system approaches that apply a response to provoked emotions. Different interpretations and areas of research in affective computing were illustrated to provide an overview of the current approaches in the area in which the current thesis is located. Table 2.3 illustrates the main areas with major impact, such as optimal learning sys-

Affective Systems and Applications	**Recognition**	**Provocation**	**System Emotions**	**Goal**
affective learning systems	X	X		optimal learning
hyper-narrative movies	X	X		interactive entertainment
mimicry recognition	X			behaviour influence
depression recognition	X			medical prevention
virtual avatars	X	X	X	behaviour modelling
physical androids / social robots	X	X	X	behaviour modelling
dialogue systems	X	X	(X)	human-computer interaction
biofeedback systems	(X)	X		behaviour influence
affective games	X	X		interactive emotional experience

Table 2.3.: Affective Systems and Applications Overview

tems, which are able to adapt to emotional reactions. The column system emotions shows systems that models emotions for the system itself like for instance virtual avatars.

2.4. Emotion Recognition

Reliable emotion recognition is crucial for affect-sensitive interfaces, as an adequate reaction to users' emotional states is required. The recognition is a tedious task, as emotions are constructs with fuzzy boundaries and individual variations in expression and experience (Calvo and D'Mello 2010). Related research is often based on multimodal approaches, because every detection modality has advantages and disadvantages. Emotions can be recognised by, for instance, evaluating facial expressions, physiological data, voices, spoken words, written text, thermal changes, gestures or posture. Even mouse activity has been evaluated to classify affective states, with features such as pressure or frequency of clicking and trajectories of movement (Yamauchi 2013).

Speech based emotion recognition has often been applied in affective computing research (Snel and Cullen 2013; Li et al. 2013; Giraud et al. 2013b; Yu, Zhou, and Riekki 2009). Voice is a promising signal for affect recognition, due to its low-costs, non-intrusiveness and fast time resolution. Emotion recognition from speech analysis can on the one hand be done by evaluation of the semantic of spoken words, and on the other hand by a prosody analysis. Pitch can be matched to arousal, but the recognition rate of basic emotions is somewhat lower than that of facial expressions (Calvo and D'Mello 2010). Sapru and Bourlard (2013) found that language style and acoustic features are correlated with social roles, which makes emotion recognition from speech a challenging task.

Text-based emotion recognition was carried out by Chatzakou et al. (2013) who analysed Twitter feeds regarding their emotional intensity, or by Barros, Rodriguez, and Ortigosa (2013) who analysed poems. A main challenge of a text-based emotion recognition is disambiguation.

Another important part of human communication that has been studied for affect recognition is **laughter detection** (Urbain, Çakmak, and Dutoit 2013). Petridis, Leveque, and Pantic (2013) for instance developed a system to automatically detect laughter from audio and video features. Moreover, McKeown et al. (2013) and Griffin et al. (2013) used body movement to recognise and categorise laughter.

According to Calvo and D'Mello (2010) **postural and body language** based emotion recog-

Modal	Recognition System
speech	speech recognition
speech / body	laughter detection
hand and arm gestures	gesture recognition
body posture	posture recognition
facial posture	facial expression recognition
body signals	physiological signal interpretation

Table 2.4.: Emotion Recognition Overview of Modalities

nition is very promising for affect recognition. Giraud et al. (2013a) found that displacements of the centre of gravity during a public speaking task are related to negative emotions. They analysed video data combined with the measurement of a force plate. In their setup they did not analyse the gestures of a speaker, although it has been shown that **Gestures** are related to affective changes (Samadani et al. 2013; Chellali and Hennig 2013). Table 2.4 provides an overview of emotion recognition systems.

The most promising techniques for emotion recognition are based on **facial Expressions** and **physiological sensors**. Both have been integrated into the EmotionBike system, and more details are provided in the following. Furthermore, an overview of multimodal affective computing databases is presented and fusion strategies are described, as many emotion recognition systems are based on multimodal input.

2.4.1. Facial Expression Recognition

Facial Expression Recognition (FER) has been studied in various applications (Almaev and Valstar 2013): public speaking tasks (Giraud et al. 2013b), affective cinemas (Abadi et al. 2013), or in measuring political preference for an election debate (McDuff et al. 2013). However, the majority of available commercial software has a focus on market research. In the following, current challenges of FER will be presented, and two state-of-the-art software tools will be described. However, automatic FER systems do not match humans (Asthana et al. 2009; Brick, Hunter, and Cohn 2009; Hoque, El Kaliouby, and Picard 2009).

Before a facial expression analysis can be performed, the face needs to be detected and the image normalised in size and rotation (Garbas et al. 2013). This rotation and normalisation is a challenging task for the facial expression recognition tools, which are available on the market. In addition, changes in lighting conditions and clutter (Zeng et al. 2009) are difficult, as is too much head movement (Dhall 2013). Moreover, context control is important for facial expression recognition (Aviezer et al. 2008; Barrett et al. 2007). Further information about current challenges as well as a comparison of tools available on the market is presented in our work published in Bernin et al. (2017).

Facial expression recognition systems that are based on machine learning (Donato et al. 1999; Bartlett et al. 1999) often apply FACS (see Section 2.1.1) coded images as training input. Different areas of interest in the face are categorised as AUs and can be linked to the basic emotions. According to Bartlett et al. (2003), spontaneously occurring facial expressions differs substantially from acted ones, however many labelled databases are based on acted facial expressions. A reason for this is that the labelling of spontaneous facial expressions is very time consuming and error prone (Zeng et al. 2009). Thus, many available FER systems utilise artificially portrayed data under controlled recording conditions as training data (Garbas et al. 2013).

A tool that is commonly used in related literature (Stratou et al. 2013; Grafsgaard et al. 2013) is called **CERT** (Computer Expression Recognition Toolbox) and was introduced by Littlewort et al. (2011). It is based on machine learning algorithms to train the emotion recognition in image sequences showing FACS coded emotions. Gabor filters, the AdaBoost algorithm and support vector machines are utilised for classification (Littlewort et al. 2011; Littlewort et al. 2006; Bartlett et al. 2003; Bartlett et al. 2006). It provides a probability value for the successful recognition the emotions: joy, disgust, anger, fear, neutral, sad, surprise and contempt. In this PhD thesis CERT was applied to the first experiment (see Chapter 4) but was substituted with its successor called Emotient[4], due to its real-time abilities (see Chapter 5).

As already mentioned, the emotion theories are controversial, thus facial expression recognition has also utilised facial feature points instead of calling it AUs (Hakim, Marsland, and Guesgen 2013; Bilakhia, Petridis, and Pantic 2013; Petridis, Leveque, and Pantic 2013).

Another library that is commonly used in related literature is called **SHORE** (Sophisticated High Speed Object Recognition Engine) and was introduced by Küblbeck and Ernst (2006). They presented an illumination invariant approach to face detection that was enhanced by gender and facial expression classification by Ernst, Ruf, and Kueblbeck (2009). SHORE detecs the emotions: happy, sad, surprised and angry. Garbas et al. (2013) enhanced the system by a continuous facial expression recognition system to measure valence in real-time. The author applied SHORE in my previous works with textile designers (Müller et al. 2012b; Müller et al. 2012a; Müller 2013), and in related research SHORE has been utilised to for instance recognise emotions in the interaction between children with autism and an android (Mazzei et al. 2012), for basic emotion analysis (Stratou et al. 2013) and even for FER of great apes (Ernst and Kublbeck 2011). However, in this PhD thesis CERT and its successor Emotient are applied due to more available emotions. In the EmotionBike experiments frustration is provoked and therefore needs to be recognised.

[4] https://imotions.com/emotient/

Modal	Recognition System	Classification
speech	speech recognition	
speech / body	laughter detection	
hand and arm gestures	gesture recognition	
body posture	posture recognition	
body signals	physiological signal interpretation	
facial posture	facial expression recognition	facial feature points
		facial action coding system
		direct emotion mapping

Table 2.5.: Emotion Recognition Overview: Facial Expression Recognition

Table 2.5 summarises the ways of classification in facial expression recognition systems. A classification of emotions by facial feature point evaluation is as acceptable as applying FACS to the recognition system or mapping the emotions directly. All methods of classification are based on machine learning techniques, thus the training data are more crucial to the results than the classification.

2.4.2. Physiological Data Analysis for Emotion Recognition

Related research in the field of affective computing often integrates physiological data analysis, which allows to observe reactions of the autonomic nervous system (ANS) (Raaijmakers et al. 2013). An early study of correlations between physiological reactions and basic emotions was presented by Ekman, Levenson, and Friesen (1983). These correlations have also been investigated in more recent works, for instance by Jang et al. (2011). Physiological data have been utilised to evaluate affective interactions. Prendinger, Becker, and Ishizuka (2006) evaluated affective interactions with a physiology-based approach, instead of, or combined with (Becker et al. 2005), utilising self-assessment questionnaires.

Physiological data analysis has also been applied for fitness measurement systems. People who are focused on improving their health through sports need to measure and monitor their physical fitness, especially if they want to prevent obesity. Therefore Shih, Chao, and Hsu (2011) introduced a health management system to support medical diagnoses and to quantify training based on physiological data.

In affective computing research that applies physiological data, many terms are used synonymously. Electrodermal Activity (EDA) for instance is also called Skin Conductance (SC). Furthermore, many abbreviations differ between publications. Therefore an overview of applied terms and their abbreviations is presented.

Physiological data analysis has been used to analyse mental and physical user performance to enhance affective computer systems (Akhanda, Islam, and Rahman 2014). There have been many attempts to collect the data in a non-invasive way (Lee et al. 2010). Liu, Wang, and Wang (2013), for instance, by recognising pulse rates with a standard web cam. A commonly studied part of the physiological data is the **pulse**. A **Photoplethysmogram (PPG)** sensor is a commonly used technique (Jang et al. 2011; Lee et al. 2010), mostly based on a pulse oximeter that illuminates the skin and measures changes in light absorption. Other terms in affective computing are **Skin Blood Flow (SBF)** (Giraud et al. 2013b) or **Pulse Rate (PR)** (Liu, Wang, and Wang 2013).

Skin conductance (SC) has been shown to be a reliable source for affect recognition (Cheng and Liu 2013; Jang et al. 2011; Chellali and Hennig 2013; Parnandi, Son, and Gutierrez-Osuna 2013; Nacke et al. 2011; Giraud et al. 2013b; Abadi et al. 2013; Koelstra et al. 2012; Negini, Mandryk, and Stanley 2014; Healey and Picard 2005). The terms **electrodermal**

activity (EDA) and **galvanic skin response (GSR)** are used synonymously in related research.

A promising source for emotion recognition is **body temperature**. Thermal cameras have been used to analyse thermal infrared images to measure facial expressions (He et al. 2013). Sharma et al. (2013) analysed thermal and video data from participants' faces for automatic stress detection. **Skin temperature** has also been utilised for affect recognition (Nacke et al. 2011; Koelstra et al. 2012). The abbreviation used for the term skin temperature differs between works; Jang et al. (2011) uses (SKT) and Giraud et al. (2013b) apply (ST).

Stress has been shown to increase **blood volume pressure (BVP)** (Koelstra et al. 2012), or **blood pressure (BP)**. In addition, BP can be used to calculate the **heart rate (HR)** by identification of heart beats, periods between the heart beats and the **heart rate variability (HRV)** (Koelstra et al. 2012). BP has been analysed in various related studies (Singh, Singla, and Jha 2009; Nacke et al. 2011; Shih, Chao, and Hsu 2011; Munia et al. 2012; Sharma and Gedeon 2013) and according to Koelstra et al. (2012) BP and HRV correlate with emotions due to an increase of blood pressure in stressful situations. According to Al Khatib et al. (2006) real-time monitoring and analysis of **electrocardiography (ECG)** data in biomedical applications provide huge potential health benefits. Thus, an analysis of ECG data is also common in related research (Jang et al. 2011; Singh, Singla, and Jha 2009; Munia et al. 2012; Sharma and Gedeon 2013; Akhanda, Islam, and Rahman 2014; Nacke et al. 2011; Healey and Picard 2005). **Heart rate variability (HRV)** (Schaaff and Adam 2013; Koelstra et al. 2012; Giraud et al. 2013b), **instantaneous heart rate (IHR)** (Giraud et al. 2013b) and **heart rate (HR)** (Vachiratamporn et al. 2014) have also been applied for emotion recognition.

Electroencephalography (EEG) has been claimed as very promising for emotion recognition (Calvo and D'Mello 2010). It has often been studied (Munia et al. 2012; Sharma and Gedeon 2013; Akhanda, Islam, and Rahman 2014; Abadi et al. 2013; Koelstra et al. 2012; Islam et al. 2014; Vachiratamporn et al. 2014) but requires a very strict experimental setup.

Respiration varies with different emotional states (Koelstra et al. 2012; Nacke et al. 2011). Plarre et al. (2011) found respiration features to be highly discriminatory of physiological stress. Recreation phases have been linked to slow respiration rates, and the cessation of respiration, as well as irregular or rapidly varying respiration can be interpreted as being caused by highly aroused emotions such as anger or fear (Koelstra et al. 2012). Furthermore, it has been shown that **hormones** such as cortisol, which can be picked from salivary samples, are related to a person's stress level (Giraud et al. 2013b).

An **electrooculography (EOG)** signal can be used to recognise emotions (Akhanda, Islam, and Rahman 2014) such as anxiety. Blinking affects the EOG signal by raising detectable peaks (Koelstra et al. 2012). The **electromyography (EMG)** signal can also be used for emotion recognition (Nacke et al. 2011; Akhanda, Islam, and Rahman 2014; Healey and Picard 2005), as the activity of the zygomaticus major (a facial muscle) is activated when people laugh or smile (Koelstra et al. 2012). In addition, EMG signals can be utilised to measure the eye-blink startle response, which is related to negative emotions. A stimulus is white noise at 100-110dB, although Chittaro and Sioni (2013) found that the alternative sound of an explosion in a virtual environment has the same effect. Table 2.6 provides an overview of signals for physiological data interpretation.

For the first iteration (see Chapter 4) a certified medical system for physiological data acquisition was integrated into the EmotionBike system. The BIOPAC system[5] has a long tradition in physiological data analysis (Yu et al. 2004) and has often been used to validate developed measurement systems (Lee et al. 2010; Liu, Wang, and Wang 2013). It is commonly used in related research (Jang et al. 2011; Cheng and Liu 2013; Sareen et al. 2008; Munia et al. 2012; Akhanda, Islam, and Rahman 2014; Islam et al. 2014).

Another company providing systems for physiological data acquisition that has been used in related research (Parnandi, Son, and Gutierrez-Osuna 2013; Negini, Mandryk, and Stanley 2014; Vachiratamporn et al. 2014) is Thought Technology Ltd.[6]. Although in the second iteration, which is presented in this PhD thesis (see Chapter 5), the physiological data acquisition has been enhanced by Plux sensors[7], as they support wireless Bluetooth transmission. Furthermore, they have a broad range of sensors (TMP, EEG, BVP, EMG, EDA, ACCEL, RESP, ECG, FORCE, ANG, RSP, LUX, MAG) and are very accurate with 16bit resolution and up to 1000Hz. In addition, they provide APIs for the languages used in the distributed EmotionBike system C++, .NET, Android and Python.

2.4.3. Multimodal Affective Computing

Affect recognition has often been described as an inherently multimodal task (Wahlster 2006; Kächele et al. 2015). Multimodal systems are described as being more natural to users. This implies that the data are processed in real-time (Dumas, Lalanne, and Oviatt 2009). It has been shown that multimodality enhances the robustness of emotion recognition, as combining different modalities improves recognition results (Abadi et al. 2013). Some emotions are

[5] https://www.biopac.com/
[6] http://www.thoughttechnology.com/
[7] http://biosignalsplux.com/index.php/en/

Modal	Recognition System	Signal
speech	speech recognition	
speech / body	laughter detection	
hand and arm gestures	gesture recognition	
body posture	posture recognition	
facial posture	facial expression recognition	
body signals	physiological signal interpretation	pulse
		electrodermal activity
		body / skin temperature
		blood pressure
		heart rate
		hormones
		electrocardiography
		electroencephalography
		respiration
		electrooculography
		electromyography

Table 2.6.: Emotion Recognition Overview: Physiological Signal Interpretation

for instance easier to recognise from facial expressions and others from physiological data. An example is smiling, which is difficult to interpret from a facial expression, due to Hoque and Picard (2011)'s findings of many people smiling in cases of natural frustration. Therefore related works are often based on multimodal sensory input (Koelstra et al. 2012; Abadi et al. 2013; Giraud et al. 2013b).

Many related studies collect their own data, due to the often-criticised lack of **databases** including spontaneous expressions of affective states (Giraud et al. 2013b). Sharma et al. (2013) for instance created the ANU Stress database called ANU StressDB, consisting of facial video and thermal data from 35 participants watching six stressful and six non-stressful video film clips.

However, at least some validated facial expression recognition databases are available to test or improve recognition algorithms. Almaev and Valstar (2013) utilised the databases MMI (Pantic et al. 2005; Valstar and Pantic 2010) and Cohn-Canade (Kanade, Cohn, and Tian 2000). The MMI database was conceived by Maja Pantic, Michel Valstar and Ioannis Patras. Petridis, Leveque, and Pantic (2013) applied audio and visual features from the SEMAINE (Sustained Emotionally coloured Machine- human Interaction using Nonverbal Expression) database (McKeown et al. 2012). Bilakhia, Petridis, and Pantic (2013) used data from the MAHNOB (Multimodal Analysis of Human Nonverbal Behaviour in Real-World Settings) database (Sun et al. 2011; Petridis, Martinez, and Pantic 2013). More details about databases and benchmarking of FER algorithms can be found in our work published in Bernin et al. (2017).

Data annotation is a challenging task when setting up an affective computing database. According to Gunes et al. (2011b) there is no universal coding scheme for the annotation of affective data that is accepted overall. Furthermore, classifying facial actions manually is very time consuming (Donato et al. 1999; Calvo and D'Mello 2010). Yu, Zhou, and Riekki (2009) use a free video annotation programme to self-annotate emotions in video clips, which is called "Anvil" and was introduced by Kipp (2001), alhough Yu, Zhou, and Riekki (2009) applied a semi-automated emotion recognition from facial expressions and speech.

A manual FACS annotation is very labour intensive. Therefore many tools have been developed with a focus on automated FACS coding, which can easily deal with a higher sample size (Grafsgaard et al. 2013).

In their work, Constantine and Hajj (2012) claimed the collection of a ground-truth emotion set to be a huge problem, as defining one is a challenging task. However, such a data set is essential for the training of classification models. A ground-truth could be based on automatic recognition, self-assessment or observer-assessment, however all possibilities advantages and disadvantages. All three have been applied in the thesis at hand.

2.4.4. Fusion Strategies of Multimodal Affective Computing Systems

According to Abadi et al. (2013), Lee and Park (2008) and Koelstra et al. (2012) two main fusion strategies exist. Abadi et al. (2013) described an (1) early integration and a (2) late integration which can also be called (1) feature fusion and (2) decision fusion. In (1) a single vector is produced that contains all features from different sensors and is used as an input to a classifier, while in (2) each sensor produces its own feature vectors, which are fed to different classifiers; the results are combined at the end. The (2) late integration is preferred (Koelstra et al. 2012; Abadi et al. 2013; Sapru and Bourlard 2013; Petridis, Leveque, and Pantic 2013), due to the possibility of setting up an optimal weighted method to adjust the modalities, and the fact that it allows to flexibly model each sensor's asynchronous characteristics. Koelstra et al. (2012) mention that the reliability of the channel can easily be taken into account by making the final decision accordingly. However, a feature-based approach (1) supports a straightforward implementation and allows to consider synchronous characteristics of each sensor.

Dumas, Lalanne, and Oviatt (2009) mention three types of fusion: data fusion (3), feature fusion (1) and decision fusion (2). Data fusion is based on raw data, while feature fusion combines previously extracted features of each sensor and is missing in the previous classification. Data fusion is used in case of similar modalities, like for instance two cameras recording the same scene. Feature fusion is used in case of time synchronised modalities such as speech and lip movement. They also described decision fusion as the common type of fusion in multimodal systems, due to its ability to fuse loosely coupled modalities and thus its high resistance to noise and failures. Gunes et al. (2011b) also mention three types of categories: feature level fusion (1), event level fusion (4) and decision level (1) fusion, but they do not describe event level fusion further.

Many related emotion recognition studies utilise **machine learning** techniques to classify emotions. Jang et al. (2011) apply Naïve Bayes and Support Vector Machines (SVM) among others to classify happiness, sadness, anger, fear, disgust, surprise and stress by physiological data analysis (ECG, EDA, PPG, SKT). Cheng and Liu (2013) use an SVM to classify happiness, sadness and fear by physiological data analysis (SC), and Sharma and Gedeon (2013) utilises SVMs and neural networks for stress recognition.

2.4.5. Emotion Recognition Summary

In the field of affective computing it is important to provoke emotions that are natural and not acted, as they differ from acted emotions. Many related studies set up their own database, which includes spontaneous emotional expressions. This facilitates more robust emotion recognition techniques in realistic environments. Furthermore, a multimodal sensory input has shown to enhance emotion recognition. Commonly used sources for emotion recognition in related works are facial expression recognition and physiological data analysis. The most common fusion strategy is late integration or decision level fusion, as they allow to flexibly model each sensor's asynchronous characteristics.

Emotion recognition in controlled laboratory settings shows high detection rates. Various approaches have been developed during the last years, and some of them are quite sophisticated, while others are not. There is definitely more research needed to improve the available approaches. Modern techniques such as for instance deep learning might achieve more promising results in the future. Today they are mostly based on basic emotions and often lack a more differentiated consideration.

2.5. Experimental Methodology

According to Jürgen and Döring (2006) **quantitative research** applies statistical methods to analyse measured data. **Qualitative research** methods are for instance interviews, introspection, observation or think-aloud methods (Mey and Mruck 2010). Furthermore they are focused on an interpretation of collected data (Jürgen and Döring 2006). **Mixed methods** can be described as a combination of quantitative and qualitative research methods (Mey and Mruck 2010) and is a commonly used practical research approach (Jürgen and Döring 2006; Kuckartz 2014). In this PhD thesis qualitative and quantitative methods are combined to provide the contributions to knowledge defined in Section 6.2.

The definition of the term **experiment** differs between sciences. An experiment in physics has other requirements than an observational study in human sciences or an experiment in psychology or medical sciences. This work is written by a computer scientist and it answers research questions from the field of HCI, therefore reduced, but also reproducible, experiments are required instead of longitudinal case studies. In HCI applications an extensive interaction can be controlled in a defined experiment. This goes beyond a classical observational study to meet the requirements of quantitative analysis. A process from the field of experimental psychology has been applied in this work, as illustrated in Figure 2.5. It is common practice in affective computing research to conduct small and short case studies to evaluate affective systems, as presented in Section 2.5.1.

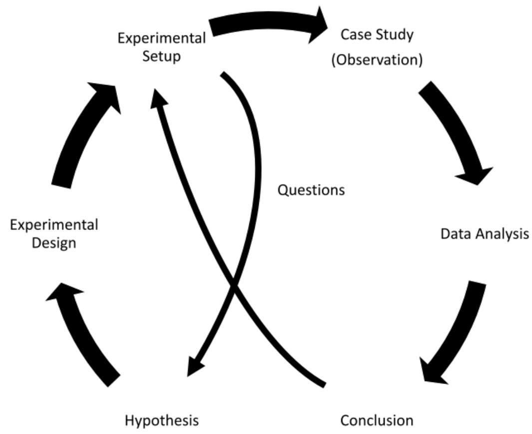


Figure 2.5.: Process of Forming Hypothesis in Experimental Psychology (Sarris and Reiß 2005)

To summarise, affective computing research is often highly interdisciplinary, and it is quite common to mix the various experimental strategies. Thus, a selected overview of affective computing study designs is presented in the following.

2.5.1. Related Affective Computing Study Designs

Ethical approval is essential to conduct an affective computing study. Related study designs also have in common that an **experimental procedure** is described, which often contains training, base-lining, emotional provocation and data collection for emotion recognition. This includes the **assessment of emotions** or data annotation. Furthermore, the **participants' profiles** and the **system setup**, including the sensor placement, is described. Additional parts of some works are a **data analysis** or an adequate system response.

It has been claimed in related research that it is difficult to transpose results from laboratory experiments to spontaneous emotional experiences (Kaiser and Wehrle 1996). Therefore, the EmotionBike project was located in a furnished smart home environment in the second experiment (see Chapter 5), to avoid feelings of discomfort.

The work of Giraud et al. (2013b) describes a related study design. They recorded speech, videos of facial expressions and body movements, balancing on a force plate, and physiological measures with a multimodal system setup. Their work describes a protocol for stress provocation in a public speaking task. This is a modified version of the widely used Trier Social Stress Test (Kirschbaum, Pirke, and Hellhammer 1993), which is known to induce stress in laboratory settings. Their work includes individual profiles based on Big Five personality traits (John and Srivastava 1999) (see Figure 2.6). In addition, Giraud et al. (2013b) used questionnaires to assert emotional states, personality profiles and relevant coping behaviours to study how participants cope with stressful situations. They collected their data in a database.

The EmotionBike is a multimodal system, which is further described in Chapter 3.2. In the experiments described in this PhD thesis, stress is induced by the presence of an observer. In distinction to other works the EmotionBike system is an exergame, which allows to evaluate the effects of physical effort in affective systems.

Participants' Profile

Persons with medical illnesses or on medication influencing the central nervous system (Jang et al. 2011) were not allowed to participate, in order to guarantee valid measurements. Before the experiment started the participants had to fill out a questionnaire to ensure these requirements. Moreover, it has shown that it is important to consider a participant's **game experience** in affective systems (Wang and Marsella 2006). Related research is often based on students as participants (Giraud et al. 2013b). In the presented experiments a more varied participant profile has been applied.

Personal fitness aspects might influence the results of an affective exergame. Hoda, Alattas, and El Saddik (2013) categorised two groups: persons who do no sports and persons who exercise 60 to 120 minutes per week.

The expression and perception of emotions is individual and differs in intensity (Calvo and D'Mello 2010). The personal information about the participants in related research often includes a categorisation in **personality traits** in order to divide them into smaller groups and reduce individual differences in the analysis, although the results are controversial. Dennis, Masthoff, and Mellish (2013) evaluated the influence of conscientiousness to emotional supportive feedback of learning effects. They found that the grade has a higher influence, although learners' conscientiousness influences the amount of advice given. In contrast, Brouwer et al. (2013)'s results indicate a negative correlation between neuroticism and skin conductance. They also found links between heart rate and skin conductance with personality. Individual physiological responses to stress have also been shown to be correlated with personality traits. Schneider et al. (2012) investigated resilience by using the appraisal ratio and examining the unique influence of personality on stress responses across multiple stressor outcomes, including affect and performance. They found that extraversion plays a role in the stress process. It uniquely predicted higher positive affect and lower negative affect in their study with arithmetic tasks as stimuli.

The most commonly used questionnaire to model personality traits is the **Big Five Trait Taxonomy** (Giraud et al. 2013b), which was introduced by John and Srivastava (1999). It allows to model the personality traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism, shown in Figure 2.6.

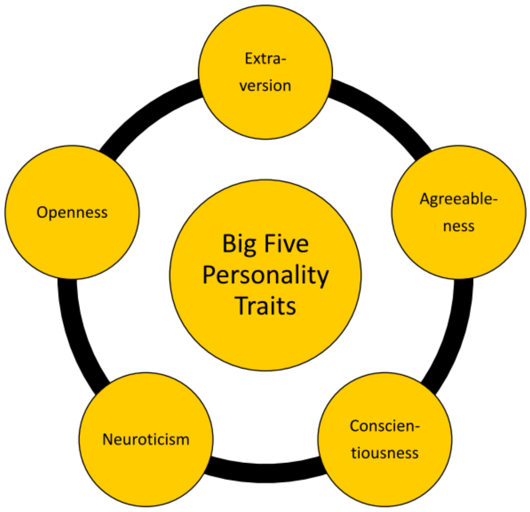


Figure 2.6.: The Big Five Personality Traits (John and Srivastava 1999)

A brief questionnaire to measure personality traits that is often used in related research (McKeown et al. 2013) was designed by Gosling, Rentfrow, and Swann (2003). The questionnaire, called TIPI, is designed for time limited research. It consists of ten questions with a rating scale from 1 ("Disagree strongly") to 7 ("Agree strongly"). The test is described as acceptable in cases of personality not being the primary topic of interest, and diminished psychometric properties can be tolerated.

For this work the author decided to apply a more detailed version of a Big Five measurement, which was designed by Satow (2012), because this questionnaire was written in German to support fluent completion.

The **number of participants** is an often-discussed problem in the experimental design of

0 - 10	(Munia et al. 2012), (Wang and Marsella 2006), (Velloso, Bulling, and Gellersen 2013), (Chellali and Hennig 2013), (Nacke et al. 2011)
11 - 20	(Akhanda, Islam, and Rahman 2014), (Giraud et al. 2013b), (Abadi et al. 2013), (Hoda, Alattas, and El Saddik 2013), (Parnandi, Son, and Gutierrez-Osuna 2013), (Vachiratamporn et al. 2014), Negini, Mandryk, and Stanley (2014), (Walmink, Wilde, and Mueller 2014), (Jang et al. 2011), (Singh, Singla, and Jha 2009), Hong, Ramos, and Dey (2012)
21 - 30	(Sharma and Gedeon 2013), (Raaijmakers et al. 2013), (Healey and Picard 2005)
31 - 50	(Koelstra et al. 2012), (Sharma et al. 2013)
n > 50	(Schaaff and Adam 2013), (Garbas et al. 2013), (Stratou et al. 2013), (Lathia et al. 2016)

Table 2.7.: Number of Participants in Related Study Designs

affective computing studies. Lathia et al. (2016) for instance analysed data of more than 10,000 participants, as they evaluated a general hypothesis. However, in the majority of related studies in affective computing the number of participants is much lower. The requirements differ between the participating interdisciplinary research fields. Table 2.7 illustrates the varying parameters in numbers of participants.

Experimental Procedure

In the beginning the participants have to be informed about the experimental procedure and the aim of the study. They must also be informed about their right to abort the experiment at any time, as it is impossible to forecast the intensity of individual emotional reactions and personal fitness. Before the exercise can start the participants must be asked to drink to avoid dehydration (Sawka et al. 2007). In an affective game, participants have to complete a **Training Scene**, as related research defines the necessity to allow the participants to become familiar with the controls (Wang and Marsella 2006).

Various **questionnaires** have been designed and evaluated in related research of affective systems. A **fun scale** was proposed by Nacke et al. (2011) to measure how much fun a participant has perceived. It ranges from 1 (not much fun) to 5 (very fun). The perceived exertion level can be evaluated with the validated **Borg scale** introduced by Borg (2004). Abadi et al. (2013) asked their participants to self-report their level of engagement, and provided a scale with a continuous range of -10 to 10. In addition O'Brien and Toms (2010) designed a multidimensional scale to measure user engagement in the context of online-shopping, and claimed that the intensity of user engagement can be influenced by emotions.

Emotion Assessment

Self-assessment of emotions is a commonly used practice in related affective computing research (Wang and Marsella 2006; Cheng and Liu 2013; Lathia et al. 2016). Various types of rating scales can be found in literature. Jang et al. (2011) for instance used a 1 to 11-point Likert scale ranging from "not happy" to "most happy" for the experienced emotions. Another way for emotion assessment is an **observer-assessment** (Afzal and Robinson 2011). This is applied by an external human observer and by trained FACS coders, who are able to map facial expressions to emotions. However, this is a very time-consuming procedure (Zeng et al. 2009).

A physiological data analysis is often also applied, as the data are less easy to manipulate for participants (Schaaff and Adam 2013). Subjective factors can influence the self-assessments, and Garbas et al. (2013) stated that even the interviewer might influence a self-assessment. Moreover, there is nearly no chance of an acquisition of valence over time using self-reports; if a task needs to be interrupted this might influence the game experience, but only asking the participants after completion might influence the responses due to the difficulty in memorising all emotions (Schaaff and Adam 2013).

The **Positive and Negative Affect Schedule (PANAS)** is another questionnaire that is commonly used in related literature (Kindness, Mellish, and Masthoff 2013a; Lathia et al. 2016; Raaijmakers et al. 2013). It was introduced by Watson, Clark, and Tellegen (1988) and allows to measure positive and negative affect. It consists of two validated 10-item mood scales, but there is no validated German translation.

The **Geneva Appraisal Questionnaire (GAO)** (Group et al. 2002) was developed by the Geneva Emotion Research Group. It is based on the component process model of emotion by Scherer (2005) and has for instance been utilised by Wang and Marsella (2006). It allows to assess an individual's appraisal process for a specific emotional episode. The underlying research field is the appraisal theory, which itself is a part of the cognitive emotion theories described in Section 2.1.5.

According to Gunes et al. (2011b), the **Self-Assessment Manikin (SAM)** introduced by Bradley and Lang (1994) is the most widely used method for self-assessment. It displays the three variables valence, arousal and dominance. It has for instance been utilised by Koelstra et al. (2012) and Schaaff and Adam (2013), who asked the participants to self-assess their emotional state with SAM after a baseline period and after playing the game. It has been utilised in the presented PhD thesis as well, since it is the most suitable emotion assessment for the EmotionBike experiments.

2.5.2. Experimental Methodology Summary

In related research, experiments are designed to evaluate affective systems and often include case studies to analyse user perception. The approach of this PhD thesis leans on the ideas of experimental psychology. The process of defining hypotheses in experimental psychology often includes iterative cycles, as shown in Figure 2.5. Therefore this work presents two iterations and designs and conducts two experiments.

The main aspects of an experimental affective computing study design are illustrated in Figure 2.7: ethical approval, system setup, experimental procedure, participants' profile, and data analysis.

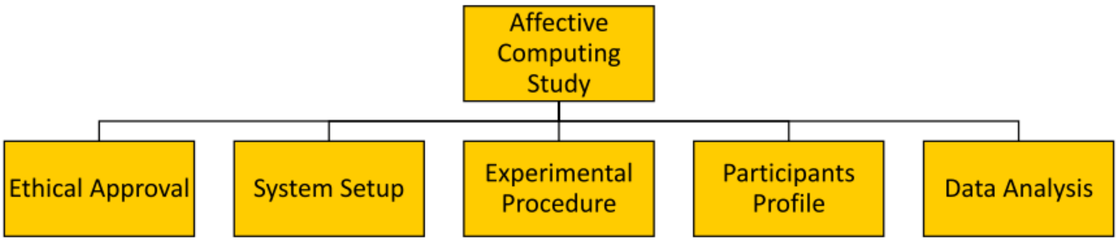


Figure 2.7.: Components of an Experimental Affective Computing Study Design

Dividing the participants into small groups with similar personality traits decrease the individual differences in the analysis, and the Big Five personality traits are commonly used in related research. The majority of affective computing studies use fewer than 20 participants.

The use of self-assessment questionnaires is controversial in related research. Self-assessment is for instance claimed to interrupt the flow, or to be influenced by personal factors (Schaaff and Adam 2013). Physiological data analyses are preferred, as they are less easy to manipulate for the participants (Schaaff and Adam 2013).

2.6. Summary of Related Work and Discussion of Open Research Issues

This thesis provides guidance for a researcher wanting to conduct an affective computing experiment in the highly interdisciplinary field. It relates to designing, building and evaluating affective systems, as presented in Figure 2.1.

2.6.1. Summary

The design of an affective system requires a suitable emotion theory. The psychological theory of basic emotions has been shown to be very promising for facial expression recognition, and the body response theory supports physiological data analysis. Discrete modelling of emotions is equally acceptable as dimensional modelling.

For the design of an affective system, the emotional provocation needs to be defined. Exergames have shown to be a very promising way of provocation, which also enhance the system by means of a physical effort.

As the field of research is growing rapidly, an overview of recent affective systems was presented. Facial expressions and physiological data analysis were shown to be very promising for emotion recognition. As multimodality enhances emotion recognition, current fusion strategies were presented. It was described that a late integration is to prefer in affective systems.

Experiments are required to evaluate affective systems, although there is still no best practice procedure in place. Therefore, related study designs were presented. Various questionnaires were applied in affective studies, and the emotions were mostly evaluated through self-assessment.

2.6.2. Discussion of Open Research Issues

Much related research focuses on improving algorithms for emotion detection (Garbas et al. 2013). This work additionally focuses on the provocation of emotions with a crafted exergame. Furthermore, much related research focuses on the recognition of emotions (Hakim, Marsland, and Guesgen 2013). In this work, emotions are analysed and provoked by game scenes. One scene for instance is designed to provoke and analyse frustration by providing the participants with a challenging task. Related research with a focus on physiological data seldom includes movement in their acquisition (Munia et al. 2012). Hoda, Alattas, and El Saddik (2013) and Hong, Ramos, and Dey (2012) applied cycling tasks in their setup, but Hoda, Alattas, and El Saddik (2013) did not measure EDA responses, and neither work provides a dynamic system response.

Evaluating affective systems requires case studies to analyse the emotional reactions of participants. Therefore many works includes user studies (Abadi et al. 2013; Wang and Marsella 2006; Jang et al. 2011; Hoda, Alattas, and El Saddik 2013), but only few collected the data in a database, because it is a tedious task. Nevertheless, some works did create databases containing non-acted emotional reactions (Koelstra et al. 2012; Sharma et al. 2013; Stratou et al. 2013), but in the presented work the non-acted emotional reactions were provoked in an exergaming context, which requires physical effort and movement.

The approach of this thesis is novel in the field of affective computing. It encompasses the recognition of emotions in an cycling exergaming setup by analysing EDA and facial expressions, where the context also includes the emotional provocation, the collection of the data in a database, and the adaptation of the game play as a response to the emotional reactions. Table 2.8 illustrates more details of selected related research in comparison to the work presented in this PhD thesis.

The next chapter presents an emotion provoking exergame system that facilitates affective computing studies. There is no realistic testbed for methods that provoke and detect emotions of users while exercising. Thus, the author needed a practical testbed to introduce new methods to the field of affective computing and to provoke real emotions. The system setup needed to be highly immersive, and not stress the player with difficult controls. Therefore, an affective exergame system was designed and set up.

Emotion recognition is a highly multimodal task; thus an open architecture was required that enabled the simple integration of various sensors to provide a reliable emotion detection. The exercising context differed from classical affective computing laboratory settings, therefore different techniques for emotion detection had to be evaluated.

An emotion provoking game was crafted to steer participants in specific emotional states. As a methodology, two experiments were designed and conducted to evaluate the emotional provocation. The first experiment was designed to evaluate single-emotion provoking game elements by emotion detection through facial expressions and physiological data. In the second experiment, multi-emotion provoking game elements were integrated and the concept of a dynamic system response was introduced. Thus, the emotion detection methods were enhanced by soft real-time abilities.

All sensor data were collected in a database of non-acted emotional reactions. The data also include personal information and the emotion assessments. Moreover, the gameplay was stored, and game events were logged to provide the possibility of testing different algorithms for emotion recognition in the future.

Current research of emotion recognition is focusing on improving recognition rates, for instance of facial expressions, in specific experimental settings. Most of the discussed methodologies of affective technologies apply to problems in classical setups. This work covers the gap of more complex and general applications. It combines different modalities and engages the user in innovative interactions. Classical approaches make it hard for researchers and practitioners to apply reliable affective computing methods to their domain. This work showcases an interactive methodology of design, prototyping and evaluation with the aid of an example in the gaming domain.

Authors	Main Category	Exergame	Emotional Provocation based on Game Events	Provocation of Stress	Discrete Modelling or Basic Emotions	Emotion Recognition based on FER	Emotion Recognition based on Physiological Data	Emotion Recognition based on EDA Data	Affective System Response	Experiment: Number of Participants	Personality Traits	Database Creation with non acted Emotions	Self-Assessment of Emotions
Koelstra et al. (2012)	DB						X	X		32		X	X
Garbas et al. (2013)	Alg					X	X	X		180			X
Hakim, Marsland, and Guesgen (2013)	Alg				X	X				0			
Munia et al. (2012)	Alg						X			10			
Stratou et al. (2013)	Study			X	X	X				53		X	
Wang and Marsella (2006)	Study		X			X				6			X
Abadi et al. (2013)	Study					(X)	X	X		15			X
Jang et al. (2011)	Study			X	X		X	X		12			X
Nogueira et al. (2016)	Study		X	X			X	X	X	24			X
Sharma et al. (2013)	DB			X		(X)				35		X	X
Hong, Ramos, and Dey (2012)	Study	(X)		X			X	X		20			X
Hoda, Alattas, and El Saddik (2013)	Study	X	(X)				X			20			X
Dennis, Masthoff, and Mellish (2013)	Alg									242	X		
Giraud et al. (2013b)	Study			X		X	X	X		19	X	X	X
Cardona et al. (2016)	Study	X	X	X			X	X		17			X
EmotionBike	X	X	X	X	X	X	X	X	X	25	X	X	X

Table 2.8.: Classification of Selected Related Research

3. EmotionBike System Design and Setup

During the last decades, the influence of affective computing has grown. Various industrial sectors evaluate users' emotions to improve marketing strategies or analyse users' attention, such as the automotive industry. The field of human-computer interaction has long apprehended the importance of emotions in communication, and various video games integrate user emotions in their game play. Emotions can be highly motivating, and a sector that benefits from increased motivation is the sports industry.

Major gaps have been identified in the related research, and these provided the motivation to build the affective exergaming system presented in this thesis. In related research there is no testbed that supports the easy evaluation of emotion recognition and emotion provocation technologies in an exercising context that also enables to conduct experiments easily. Nor is there a cycling exergame setup that is able to provoke and detect emotions and provide a dynamic game play based on a user's emotional reactions to steer the player on a predefined path of emotions.

A system is required that is able to provoke specific emotions. For this reason a game is created, as that offers many opportunities in the design of game elements. It also needs to be able to reliably detect user emotions. Thus, different emotion sensors have to be combined, which requires a distributed system setup. User studies need to be conducted to showcase the system's ability of emotion provocation and recognition.

A data acquisition system is designed to collect frontal videos and physiological data, as well as game and controller events. The emotion measurement for the presented experiment includes vision-based and physiological sensors consisting of EDA, respiration and temperature sensors.

An ergometer is enhanced by steering capabilities and software controllable pedal resistance. Ergometers are often applied for medical rehabilitation. They are also common in endurance or high intensity indoor fitness training. In fitness applications stimulation can be achieved by thrilling and exciting content. In healthcare and rehabilitation it is important to consider the individual fitness and to avoid physical or mental stress. An ergometer is an interesting setup for an affective system due to its similarity to cockpit designs. The limited monitoring area on a cycling exercise machine includes a defined position of the head and the body, which for instance simplifies the recognition of facial expressions. In addition, the subjects are located in a controlled situation to encourage context awareness and limit the degrees of freedom. In the controlled laboratory setting, the lighting can be controlled to ensure sufficient illumination of the participant's face. In this situation the attachment of physiological sensors is acceptable, as many people are used to medical stress tests on an ergometer, e.g. an exercise ECG.

This chapter presents the requirements for **Emotion Provocation** and **Emotion Detection**

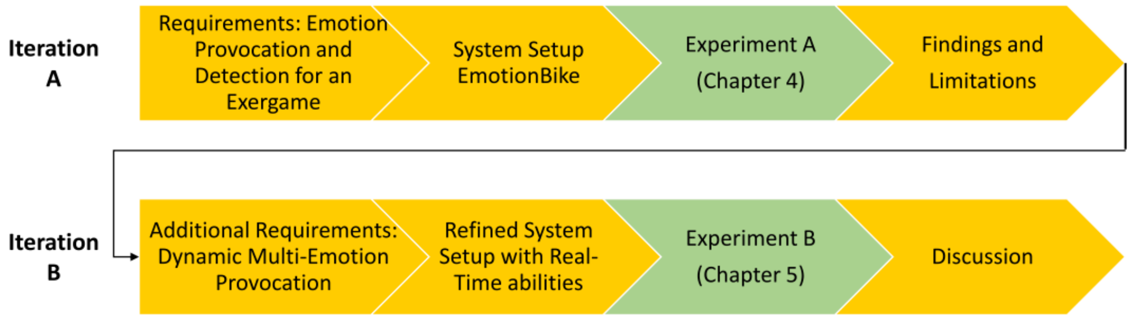


Figure 3.1.: EmotionBike: Design Process

for the exergame. Subsequently the detailed first system setup is described. As a methodology two loops are executed to improve the system. Figure 3.1 illustrates the iterative design process. The conducted case study will be described in Chapter 4. This chapter summarises the found limitations and provides refined system requirements, which focus on a dynamic system response. A second case study was also conducted and will be presented in Chapter 5. At the end, this chapter presents the conclusion of the experimental design.

3.1. System Requirements Iteration A: Emotion Provocation and Emotion Detection for an Exergame

The system requirements are illustrated in Table 3.1 and Table 3.2. The main requirement for the game was that it is able to provoke different specific emotions. The game was implemented in Unity, and the game elements were chosen from freely available assets of the Unity Asset Store. The additional game requirements were defined with the help of related literature. Affective games should be intuitive, engaging, intuitive, easy to learn, highly dynamic and enable multiple forms of adaptation (Parnandi, Son, and Gutierrez-Osuna 2013).

All game events must be logged for the emotion analysis. The collected data should enable to test different analysis methods to ensure a good quality of the results. The experimental setup requires that the scenes can be remotely started and stopped with software controls, to ensure that it is possible to conduct experiments fluently and avoid idle times.

The main requirement for the emotion detection is reliability. This includes a multimodal sensor approach as well as a self-assessment and an observer-assessment to validate the results, as described in Chapter 2. The sensor data acquisition should be minimally invasive to avoid feelings of discomfort. The use of ECG sensors, for instance, was supported by an experimenter of the same gender as the participant. An additional requirement was a defined position of the head the participants, because at the beginning of the experiment a 3D facial data analysis was planned by a team member.

The requirements for the exergame controller was the transmission of the RPM from the physical cycling exercise machine to calculate the speed in the virtual racing game. The controllable pedal resistance is important to ensure that the participant feels a difference in riding up a hill. The combination of physical effort and emotional provocation is an enhancement to other exergame systems.

The requirements for the physical setup was defined by preliminary studies. At the beginning of the project a static handlebar was used, but the users asked for steering capabilities to have a more natural experience. Many preliminary studies were made to ensure that the physical steering of the mounted rotatable handlebar feels natural to the participants. This is required to avoid discomfort, which might affect the emotional reactions to the crafted game events.

The multimodal distributed system setup requires openness to new sensors to ensure expandability. This was achieved by a standard communication protocol. The data of the sensors were collected with APIs in various programming languages. Many students were involved in the project and evaluated different emotion sensors.

The experiments needed to be reproducible, therefore all sensor and system data had to be logged. Furthermore, it was important that the system was in a clear state, even if an error occurred during the experiment. The performance was important; it had to be acceptable to the user, to avoid boredom caused for instance by extensive delays.

3.2. System Design and Setup

The experimental setup was designed to meet the presented requirements. It consists of a **Physical Exergame Controller**, **Visualisation**, **Emotion Sensors**, a **Data Acquisition System** and the **Analysis Output**. An overview of the main components is presented in Figure 3.2, and further details will be presented in this chapter. A case study was conducted with the presented system (Requirement 5.1) and will be described in Chapter 4.

1. Reliable Emotion Provocation			
1.1 Game Design (Chapter 4)	1.1.1	Different emotions	Specific emotions need to be provoked with a scene-based design
	1.1.2	Software controllable	Start of game scenes must be software controllable
	1.1.3	Logging	The game events must be logged to analyse the related sensor data
	1.1.4	Engaging	The game should be engaging to improve motivation
	1.1.5	Exciting	The game should contain exciting elements to support the fitness programme
	1.1.6	Intuitive	The game should be intuitive, the physical bicycle should be mapped to the virtual environment
	1.1.7	Easy to learn	The game and its controls should be easy to learn
	1.1.8	Highly dynamic	The game should be highly dynamic
2. Emotion Detection			
	2.1	Reliable	Different sensors need to be used to get a more reliable emotional state, additionally self- and observer- assessments are required
	2.2	Minimally invasive	The sensor data collection should be minimally invasive to make sure that the user does not feel uncomfortable
	2.3	Reasonable	The sensor acquisition should be reasonable, like exercise electrocardiograms for ergometers
	2.4	Defined head position	A defined head position facilitates an easy camera positioning to collect frontal facial video data for facial expression recognition
3. Exergame Capacity			
	3.1	Controllable pedal resistance	To provide a natural user experience, the pedal resistance must be adapted to the virtual environment. When the player cycles up a hill the pedal resistance must be increased. It is also important not to tax the user
	3.2	RPM transmission	To calculate the speed of the virtual bicycle the revolutions per minute (RPM) must be transmittable

Table 3.1.: Requirements Emotion Provocation (Part One)

4. Physical Setup			
4.1 Usability	4.1.1	Steering	Preliminary studies have shown that to provide a more natural user experience, the physical exercise machine must be enhanced by steering capabilities
	4.1.2	No emotional influence of the controls	The emotions of the user should not be influenced by too complex or difficult controls. They should be as natural as possible to avoid negative emotions
4.2 Comfort	4.2.1	Easy sensor placement	The sensors placement must be easy to avoid negative emotions
5. Experiment			
	5.1	Feasibility	A case study must be conducted to showcase the system's ability of emotion provocation and detection
6. Distributed System			
6.1 Openness	6.1.1	Communication	A standardised communication protocol is required
	6.1.2	Expandability	To improve the emotion detection the system must be expandable
6.2 Scalability	6.2.1	Loose Coupling	The system should enable a loose coupling of sensor nodes
6.3 Consistency	6.3.1	Verifiability	All sensor data must be verifiable
	6.3.2	Logging	All relevant data must be logged with the possibility to replay them
	6.3.3	Clock	The clocks of the different systems need to be synchronised and timestamps must be included in the messages
	6.3.4	Failure	The systems have to be in a controlled state even if errors occur
	6.4	Heterogeneity	The system must allow different operating systems and programming languages due to varying APIs of sensors and requirements of applied analysis tools
	6.5	Performance	The system performance must be acceptable as the user should not be bored

Table 3.2.: Requirements Emotion Provocation (Part Two)

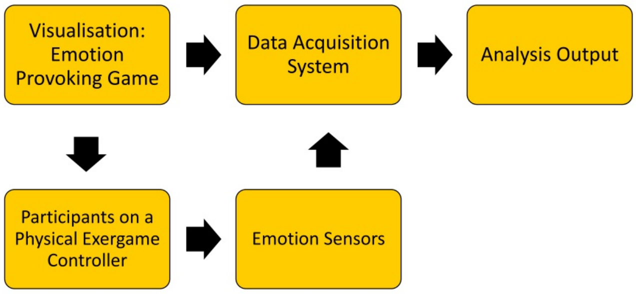


Figure 3.2.: Overview of Experimental Setup (Müller et al. 2016)

3.2.1. Physical Exergame Controller

In Section 3.1 the requirements of exergames capacities were described in detail. The most important requirement for the presented system is to provide a natural user experience and avoid any unintentional influences of the system to the emotional reactions. The physical cycling exergame controller is based on a semi-professional exercise machine, the Daum[1] premium 8i ergometer, with a maximum pedal resistance of 600 W to meet the Requirement 3 of exergame capacity and 4 a physical setup. The exercise machine operates on an embedded system and is connected with a network interface to read and control parameters such as resistance and speed, which is important the meet Requirements 3.1 and 3.2 of a controllable pedal resistance and a RPM transmission to provide a natural user experience. Therefor speed can be calculated by pedal RPM and physical pedal resistance is software controllable. The technical setup was part of a bachelor thesis by one of the students from the technical computer science department, thus further details of the technical implementation can be found in Hornschuh (2015).

The proprietary protocol structure of data packages illustrated in Figure 3.3 allows to read out for instance the calculated RPM with the command "S21" (Daum Electronic GmbH 2017).

[1] http://www.daum-electronic.de/

The calculation of the RPM value is presented in Equation 3.1; rev stands for revolution. The RPM values are transferred as speed values to the virtual game environment to move the virtual bicycle. A more physically accurate implementation was applied for the second iteration and is presented in Section 3.5.2.

SOH	Header			Data Unit			Checksum		ETB
1	2	3	4	5	...	n	n + 1	n + 2	n + 3
0x01	c1	n1	n2	Data			Checksum		0x17

Figure 3.3.: Structure of Data Packages to Communicate with the Ergometer (Daum Electronic GmbH 2017)

$$RPM = \frac{60}{\Delta t * rev} \tag{3.1}$$

The pedal resistance is calculated related to the degree of slope in the virtual game. The slope values range from 0 - 359 degrees. The range 0 - 90 degrees represents a positive slope and 270 - 359 degrees represent a negative slope. In the first experiment a rough approximation was implemented. The physical resistance was increased in steps of 10 Watt for each degree of slope, as shown in Equation 3.2. Res describes the resistance and normRes the normal resistance.

$$res = normRes + (degreeSlope * 10W) \tag{3.2}$$

The Static Handlebar was replaced by a rotatable one with the support of a project member from the mechanical engineering department. During preliminary studies participants mentioned that a rotatable handlebar is important for a natural user experience. To facilitate steering capabilities and meet Requirement 4.1.1 an incremental rotary position encoder

(Inkrementalgeber Kübler 2400[2]) was applied to measure the handlebar rotation. It was mounted at the lower end of the handlebar. A Raspberry Pi[3] was utilised to read the falling or rising edges from the encoder, and software monitors the changes of states to calculate the rotation direction. The software sends the direction data to the message broker. During tests in our project team it was found that it is important to avoid too high latencies to make sure that the user is not emotional influenced (Requirement 4.1.2). Humans are able to recognise interaction latencies above 100 ms (Bernin 2011). The measured latency of steering visualisation in the game was 80ms. Thus the cycling exercise machine can be used as a game controller, which generates the user input for the designed game environment.

3.2.2. Visualisation

The emotion provoking game was designed to meet the requirement of reliable emotion provocation (Requirement 1). The game was presented to the participants on a 42-inch flat panel monitor, shown in Figure 3.4. The user may physically accelerate and steer through the virtual game environment, as the generated controller inputs are directly transferred into the game. The simulated virtual bicycle in the game, shown in Figure 3.4, has soft real time requirements, as insufficient controls may influence the user perception and provoke frustration (Requirement 4.1.2). The game design is further described in Section 4.2; more technical details can be found in the master thesis by Zagaria (2017).

[2]https://www.kuebler.com/k2014/j/de/produkte/details/drehgeber/Rotativ/2400
[3]https://www.raspberrypi.org/

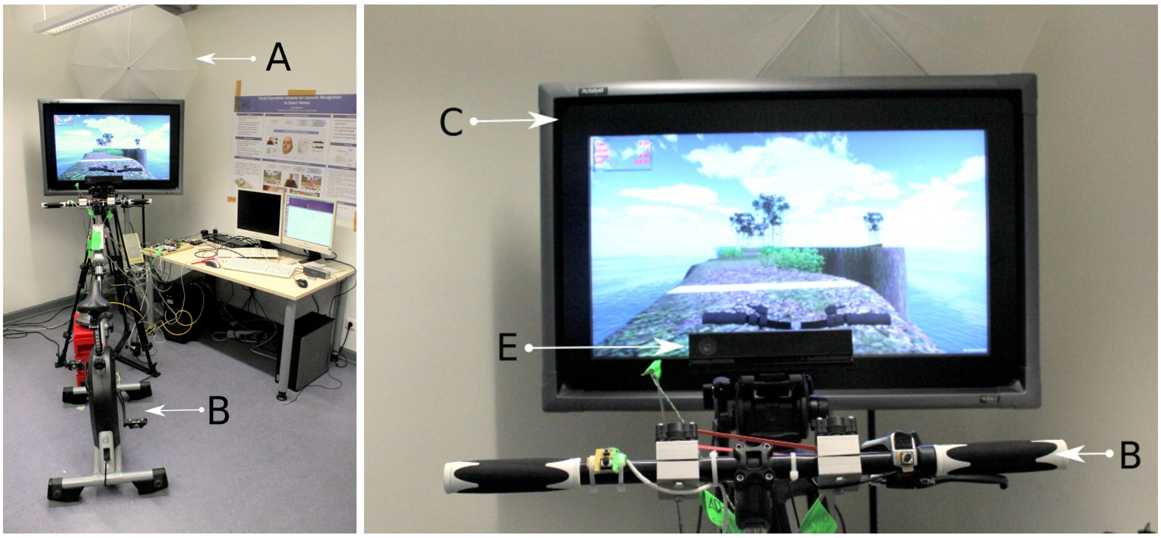


Figure 3.4.: Hardware Setup of the EmotionBike (Müller et al. 2015): Face Illumination Lamp (A), Cycling Game Controller (B), Screen (C) and Camera (D)

3.2.3. Emotion Sensors

The system setup is designed to facilitate the employment of various sensors for emotion recognition (Requirement 2). Different sensors were applied for multimodal emotion recognition to meet the requirement of reliable emotion detection (Requirement 2.1). Preliminary studies with the Emotiv Epoc[4] 14 channel neuro headset have shown that the deployment of an EEG sensor in the presented setup is not very promising. The EEG data were collected during the first experiment. Physical activity is an important part of the exergame, but movement and sweat corrupt the sensor data of an EEG. Thus, this work focuses on vision-based and physiological sensors to meet the requirement of minimally invasive detection (Requirement 2.2).

[4] https://www.emotiv.com/epoc/

Vision-based Sensors

The emotion sensor component generalises vision-based input and physiological data. A Microsoft Kinect v2 camera was placed in front of the exercise machine (Requirement 2.2) to capture video data of frontal images, as shown in 3.4. The Kinect provides HD(1080P) resolution and RGB-D(512x424) images at up to 30Hz.

Physiological Sensors

The physiological data collected during the experiment include respiration, body temperature change, and EDA. The data were provided by a physiological data acquisition system, the BIOPAC MP36. The system operates at a rate of 500 Hz and pre-processes the data with the device's internal hardware filters. These filters support gain improvement and noise reduction, and apply low-pass filtering to discard the irrelevant frequencies. In addition software filtering and smoothing of the EDA input signal is achieved by a three-level cascade consisting of a digital low-pass Butterworth filter (Rabiner and Gold 1975) (4^{th} order, cutoff frequency = 5 Hz) in conjunction with two Moving Average Filters (Smith 1997) (using "boxcar" and "parzen" kernels with a size of $375 = 0.75 \times$ sampling rate). The EDA sensor was connected to the middle and forefinger of the participants' left hand in case they stated right-handedness, according to the instructions by Boucsein (2012).

The respiration belt was placed across the participants' chest and the temperature sensor on the inner forearm (Requirement 4.2.1). Moreover, it is understandable for a user that physiological data are acquired in such a setup, as many people know exercise ECG from a medical context (Requirement 2.3).

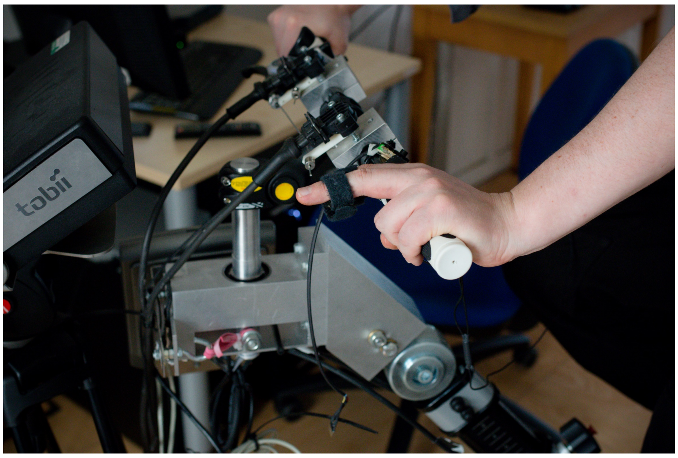


Figure 3.5.: Physiological Sensors connected to the Subject (EDA), (Müller et al. (2016), ©2016 IEEE)

3.2.4. Data Acquisition System

The different sensors use components from a distributed system of multiple client computers. All of them are loosely coupled using a message broker (Apache ActiveMQ[5]) (Requirement 6.). This blackboard architecture has been approved in previous works in our laboratory setting (Otto and Voskuhl 2010; Müller 2013). According to Dumas, Lalanne, and Oviatt (2009) all modalities in a multimodal system need to be synchronised and to have timestamps, as a framework for emotion measurement requires the ability to log and replay all events and data as described by Bernin (2012) (Requirement 6.3). In the presented system a JSON-based protocol was implemented, corresponding to the requirement of openness (Requirement 6.1). Figure 3.6 illustrates the loosely coupled system design.

Fusion:

A late integration or decision level fusion of sensor data is preferred for affective computing systems, as described in Section 2.4. It allows to flexibly model each sensor's asynchronous

[5] http://activemq.apache.org/

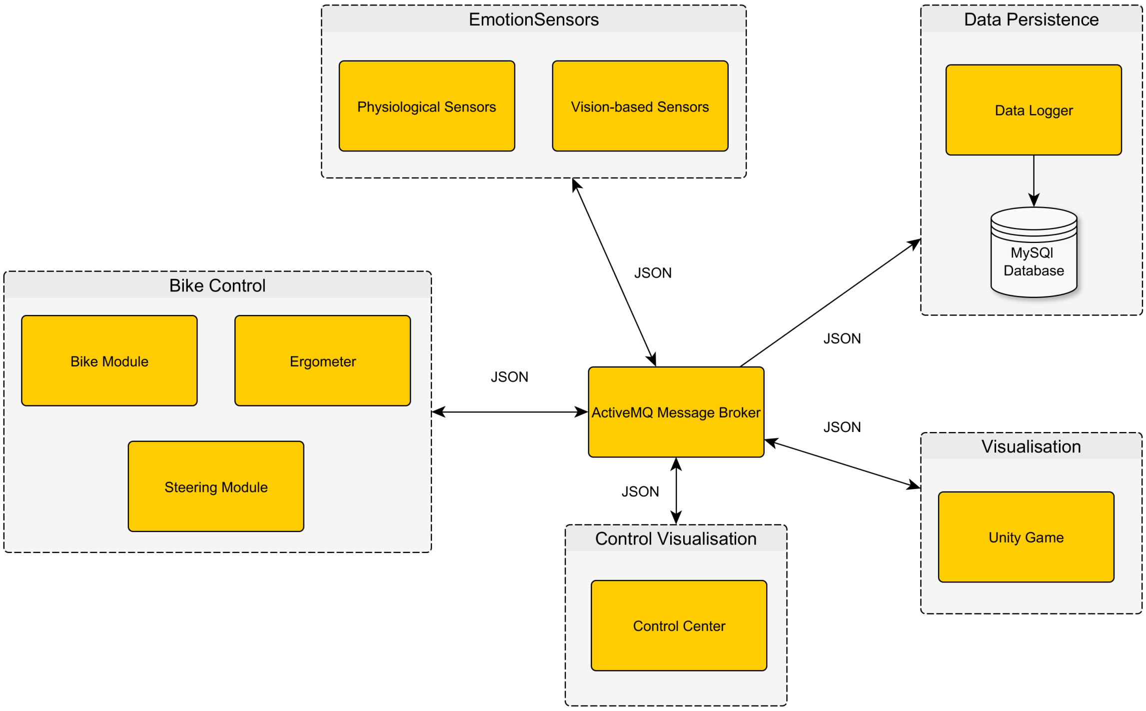


Figure 3.6.: Loosely Coupled System Design using a Blackboard Architecture

characteristics. In this work EDA sensor data are combined with facial expression recognition.

3.3. Limitations and Discussion

The system was successfully applied to a first case study to showcase the system's ability to provoke emotions, as described in Chapter 4. The system was shown to reliably detect emotions through physiological data and facial expression analysis. The results are presented in Section 4.5.

The applied CERT software for facial expression analysis does not work in real-time, but for an affective system it is important to react to users' emotions. The game design does not support a flexible adaptation, as more dynamic game scenes are required. Moreover, the mapping of the exercise machine RPMs to the virtual bicycle did not meet the participants' expectations and the majority of them stated they missed a brake and a gear shift. The laboratory environment made users feel uncomfortable, which might have influenced their emotional reactions.

Sensor data acquisition was difficult, as the cables of the applied BIOPAC system were often tangled up in the pedals. Moreover, data loss could not be easily detected as the sensor data collection was not fully monitored. The only way was to observe the visualisation of the sensor data. In one case the local storage of a computer was exceeded, which was discovered in retrospect.

The system setup is very static and does not facilitate easy user studies. However, in the field of affective computing, user and behaviour studies are very important to observe the success of the designed systems.

In summary, further enhancements were needed to provide a more natural user experience and offer a real affective system. Below the refined system requirements to overcome the illustrated limitations are presented.

3.4. System Requirements Iteration B: Dynamic Multi-Emotion Provocation

Table 3.3 and 3.4 illustrate the refined system requirements. They are based on the findings from the first case study described in Chapter 4.

The main requirement for the game was that it should be able to provoke different specific emotions, as described in Section 3.4. The aim of the second experiment was to ensure a dynamic system response that is able to adapt to user emotions. Therefore the game design must be refined to enable more dynamic and flexible control. In the first experiment all game events were static and thereby did not ensure flexibility. The difficulty of the game scenes should be configurable, as this enables the agile approach of many preliminary studies that were applied to the first experiment. Conducting the experiments is very time consuming, therefore the preliminary studies were required to ensure good experimental results.

One result of the first experiments was that the majority of the participants asked for gears and a brake shift to have a more natural user experience. Thus, both were mounted at the cycling exercise machine and integrated into the system. New game scenes were crafted to utilise these new controls.

A dynamic and adaptive game requires soft real-time analysis of user emotions. This requires the evaluation of various real-time analysis methods (details were published in Bernin et al. (2017) and are not part of this thesis).

During the refinement of the game for the second experiment and the integration of the brake, it was found that we need a more realistic physical mapping of the ergometer controls of the virtual bicycle. This had to be designed and implemented. The parametrisation of the module had to be configurable to test the module, and to meet the other requirements for the participants to feel comfortable.

The location of the ergometer had to be changed, as the room made the participants feel uncomfortable. The adjoining room was a smart home laboratory, which usually impresses people with its high level of comfort and enjoyable atmosphere. Therefore the system should was moved to the smart home.

The data acquisition of the physiological sensors caused many problems during the first experiment. The cable was tangled in the pedals, which made the participants feel very uncomfortable because they thought it was their fault. It also affected the game flow.

The system should be further enhanced to make it possible to easily conduct experiments. Different emotion sensors require different experimental designs. The EEG sensor for instance would require a static setup. Furthermore, the new game design should facilitate dynamic system control and configurable difficulties. Such setup should also allow to create a large number of experiments.

The results of the first experiment showed that more control and knowledge about the sensor states is required. The experimenter needs to be aware of the system status to ensure good experimental results. The experiments are very time consuming; therefore all sensor data should be collected and data loss should be avoided.

3.5. Refined System Setup

The system setup was refined to meet the requirements of providing a dynamic system response. Figure 3.7 illustrates the logical view of the system components. The server contained the persistence component, the experimental control, the bike module and the biosensor module. The physiological data and direction data were logged with a high frequency, thus local storage was utilised to ensure sufficient data storage capacity. The experimental

1. Reliable Emotion Provocation			
1.1 Game Design (Chapter 5)	1.1.9	Adaptation	The game must be amendable and enable multiple forms of adaptation
	1.1.10	Difficulty	The difficulty of game scenes must be configurable
	1.1.11	Dynamic	The order of the game scenes needs to be more dynamic and flexible
	1.1.12	Game Events (Chapter 5)	The control of game events needs to be software controllable and thus needs more flexibility
	1.1.13	Game Scenes (Chapter 5)	New thrilling and exciting game scenes need to be defined, which combines different emotion provoking game events to enable comparability and to use the enhanced bike controls
1.2 Social Presence	2.1.1	Provocation Intensification	The provocation needs to be increased by a socially evaluative threat. An independent observer might be perceived as a judge who evaluates the performance of the player
2. Emotion Detection			
	2.5	Real-Time	The emotion recognition must be applied in soft real time to support dynamic system control
3. Exergame Capacity			
	3.3	Gear Shift	To provide a natural user experience, a physical gear shift needs to be mounted at the ergometer and the state must be transmitted
	3.4	Brake	A physical brake has to be mounted to facilitate more thrilling game events

Table 3.3.: Additional Requirements: Dynamic System Response Part One

4. Physical Setup			
4.1 Usability	4.1.3	Physic Module	The bike module needs to be enhanced by a more realistic physical mapping of the ergometer controls to the virtual bicycle
4.2 Comfort	4.2.2	Environment	The classical laboratory setting needs to be changed to a more comfortable environment to avoid negative emotional influences
	4.2.3	Wireless Transmission	The sensor data acquisition system of the physiological data must be replaced by a wireless solution
5. Experiment			
	5.2	Substitutability	The system should provide the opportunity to change experimental setups and support to easily conduct case studies
6. Distributed System			
	6.1.2	Expandability	To improve the emotion detection the system must be expandable
6.2 Scalability	6.2.2	Configurable Monitoring	The monitoring of sensor nodes needs to be configurable
	6.2.3	Lifecycle	Sensors nodes need to implement a defined sensor lifecycle
6.3 Consistency	6.3.5	Monitoring	A monitoring component needs to be implemented to ensure that all sensor nodes are in the expected state
	6.3.5	Error Handling	The error handling needs to be improved, the sensor nodes should send messages about their state, for instance they have to check their local storage to avoid data loss

Table 3.4.: Additional Requirements: Dynamic System Response Part Two

control allowed to control the game. The application management was designed to support the easy conduct of different experiments (Requirement 5.2). It enables different configurations, e.g. determined scene orders for other experiments.

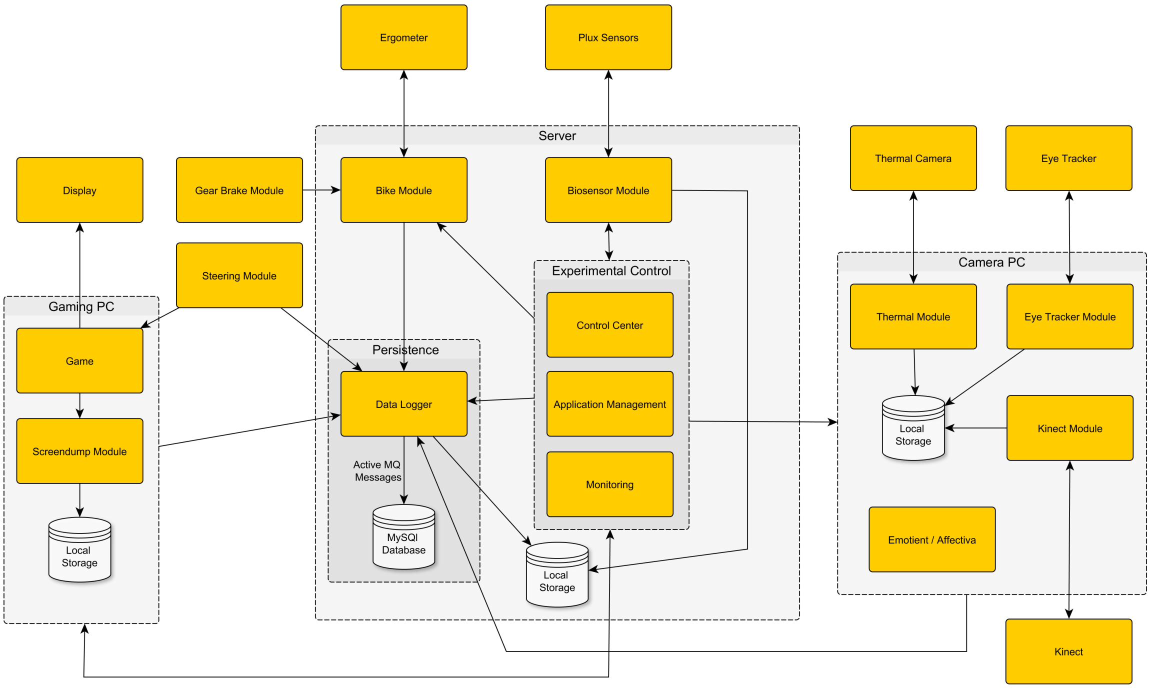


Figure 3.7.: Logical Deployment Diagram of the System Components

A gaming computer was used to ensure that the requirements were met, e.g. sufficient graphic and processor capacity to display the crafted virtual game. It contained a screendump module to allow a replay of the game activities. Another computer was applied to log and analyse the video-based sensor data.

The system architecture was enhanced by better error handling in case of failures in data acquisition. All components except the experimental control needed to implement a defined lifecycle procedure. The defined internal states of a node are: START, ONLINE, READY, MEASURE PREPARED, WORKING, WRITING, FINISHED, RESET and ERROR. An overview is presented in Figure 3.8. A configurable monitoring component was applied to ensure advanced error handling, in order to meet the requirements of scalability and consistency (Requirement 6.2 + 6.3).

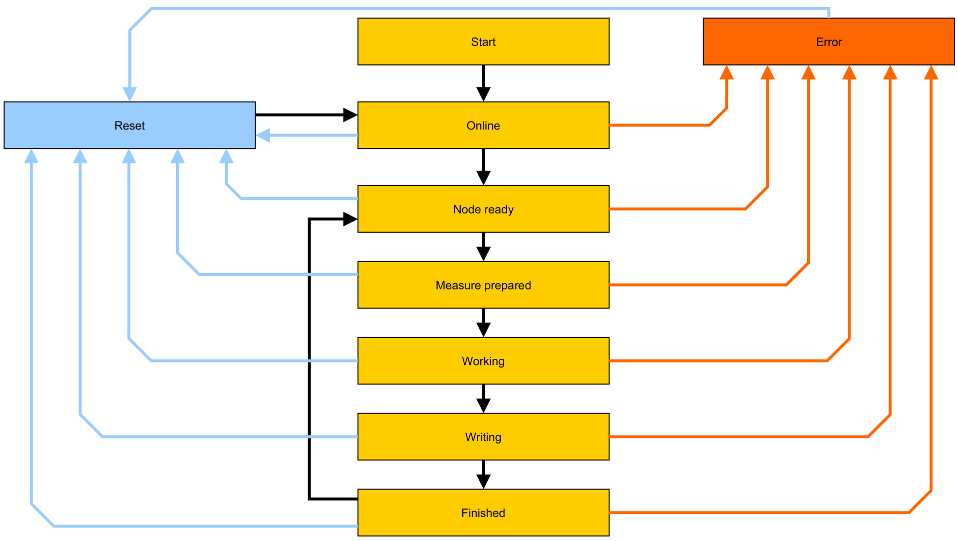


Figure 3.8.: Status of Nodes (Black: normal transition, Blue: reset triggered by the external monitoring component, Red: internal error)

The lifecycle also requires a defined procedure that has to be implemented by each node. Figure 3.9 illustrates that procedure. At first a node must establish a connection to the message broker. Subsequently each node must set up a producer and send its state changes to the monitoring channel. Furthermore, each node has to set up a producer to send heartbeat messages and a consumer must be implemented to receive control messages. In addition, node specific consumer or producer can be applied and the state changed to ONLINE. Every time the state of the node changes a message has to be sent to the message broker. In the ONLINE state, each node has time to make local preparations. The physiological data sensors for instance have to establish a Bluetooth connection and switch their state to READY after that. In this state the nodes are waiting for a control message "initializeMeasure". This message includes sender name, timestamp in milliseconds and local storage information.

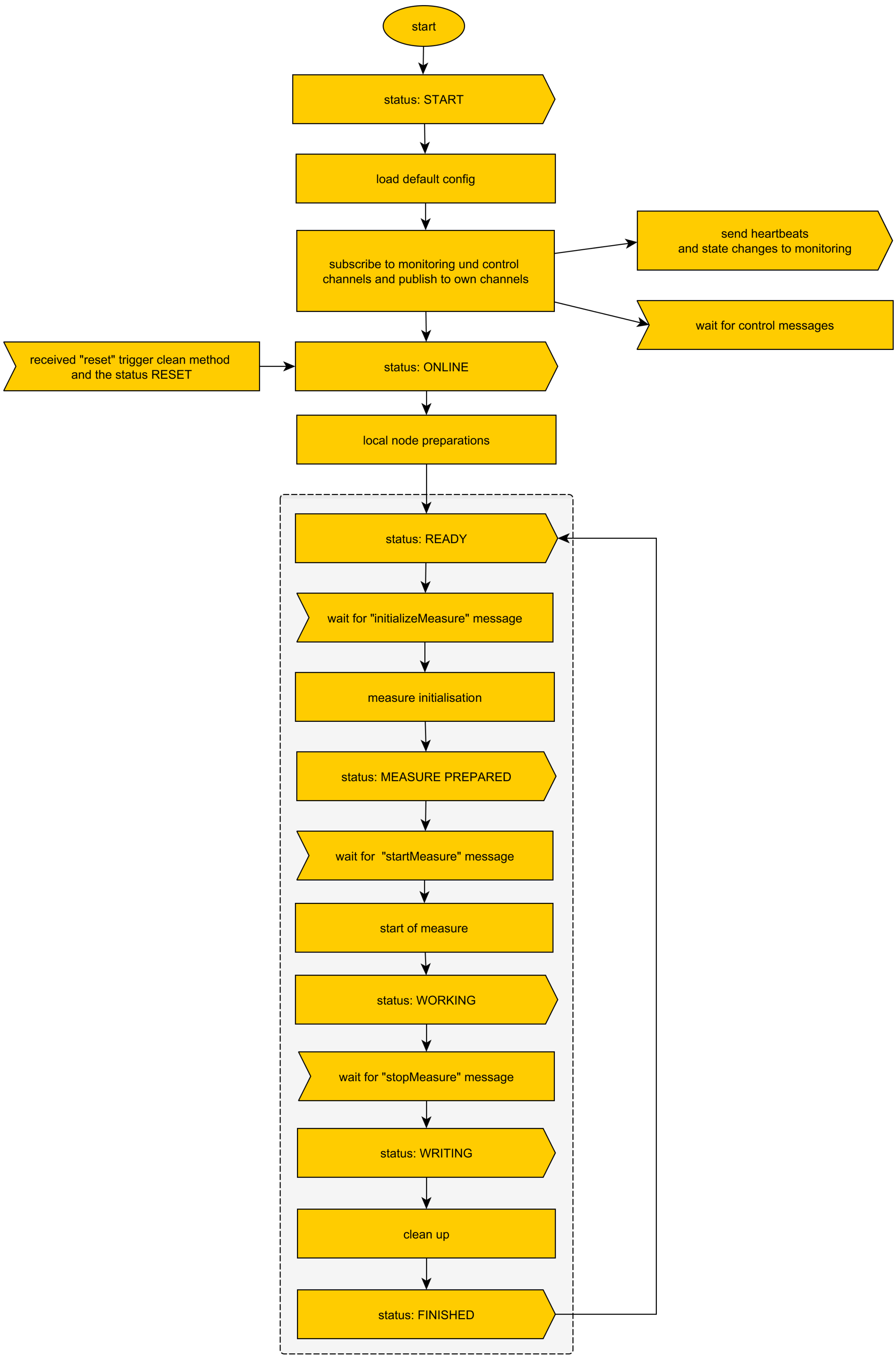


Figure 3.9.: Node Lifecycle Procedure

This storage information includes the experimental ID, the proband's ID, the session ID and the path to the server where the data should be stored. Thus a hierarchical folder structure can be applied, which makes it easy to find the data of a measure. After each node has initialised the measure the global state is switched to MEASURE PREPARED. Then the control component sends a "startMeasure" message and each node will change its state to WORKING. After the player reaches the finish line the "stopMeasure" message is sent and the nodes switch to WRITING. Subsequent to a node having finished this procedure it has to wait in the READY state for the next measure. This cycle is applied for each scene of the game except if an error occurrs, then the state is changed to ONLINE after a local RESET. When the participant finishes the experiment the lifecycle will return to the beginning.

3.5.1. Physical System Setup

The location of the EmotionBike is changed to meet Requirement 4.2.2, from a classical laboratory setting to more comfortable surroundings, the smart home called Living Place Hamburg[6]. This enhances the situational context awareness due to smart home controls. The only limitation that has to be mentioned is a missing air conditioning. The new setup is presented in Figure 3.10.

The experimental setup is enhanced by an experimental control station, as shown in Figure 3.11, to provide full system control to the experimenter. The place is separated from the participants because they should not notice that the experimenter has control over some specifically designed game elements (Requirement 4.1.12). The experimental control station includes displays for the visualisation of the applied emotion sensor data to ensure good data

[6] http://livingplace.informatik.haw-hamburg.de/blog/

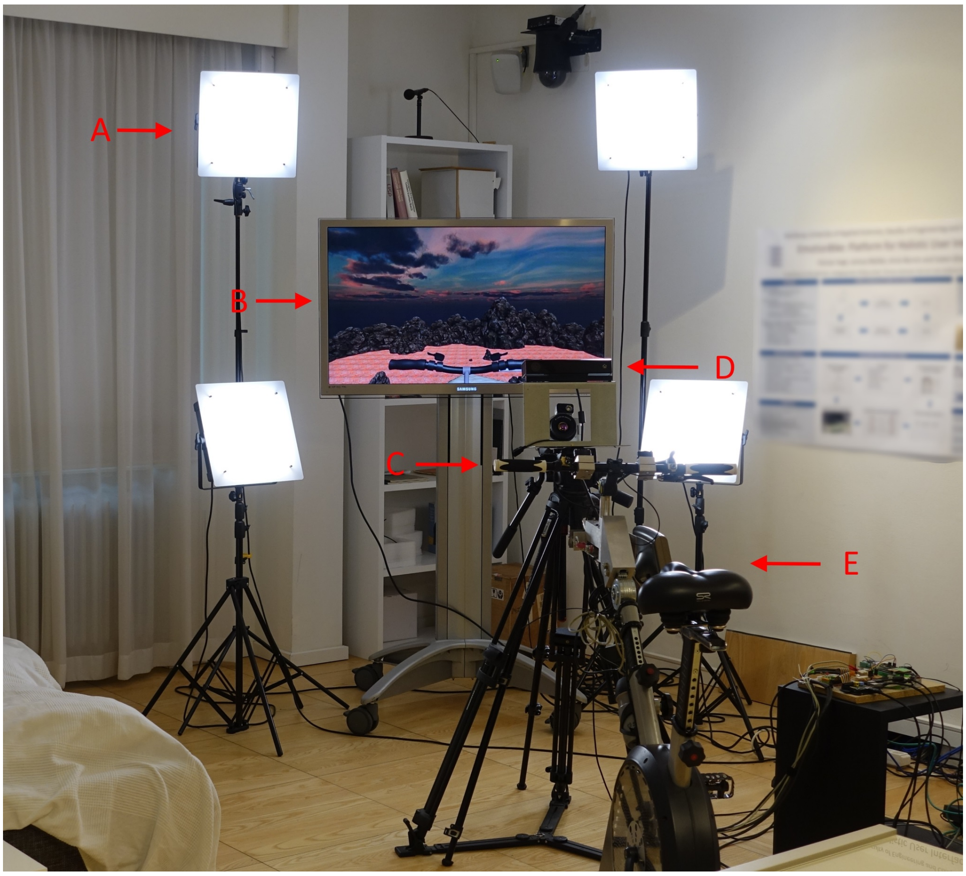


Figure 3.10.: System Setup. A: Face Illumination Lamps, B: Display, C: Rotatable Handlebar, Gearshift and Brake, D: Kinect Camera, E: Physical Exergame Controller (Müller et al. (2017a), ©2017 IEEE)

quality. In case the sensors lose their physical contacts the experimenter is able to interrupt the procedure in predetermined positions. Another display contains a mirror of the game and the video data of the participants' faces, as in some scenes the experimenter has to decide about the perfect timing of events in combination with the emotional state.

Figure 3.11.: Experimental Control Station (Müller et al. (2017b), ©2017 IEEE)

3.5.2. Physical Cycling Exergame Controller

The exergame controller was enhanced by a gearshift and a brake to improve the natural cycling experience (Requirement 3.3 + 3.4). It also offers new possibilities in the design of emotion provoking game elements. The first steps of the development were provided and described by Matthiessen (2015).

Figure 3.12 illustrates a logical view of the refined bike control modules. The bike module combines physical calculations with the states of the gear and brake to provide a more natural user experience (Requirement 4.1.3), although it must be mentioned that a physical approximation was applied that does not claim to be physically accurate. Instead the design and its implementation should provide the player with a sufficient immersive and natural feeling while driving through the virtual environment.

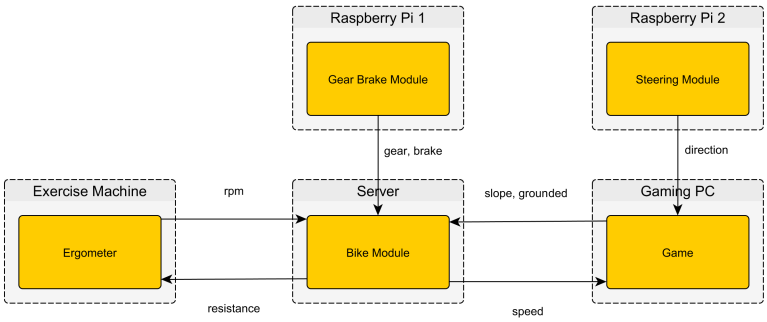


Figure 3.12.: Logical View Bike Controls

In the previous version there was no gear or brake, and the speed calculation was made in the virtual game. In the refined system the bike module calculates the speed by considering physics, RPM, the virtual environment, as well as the use of the gears and the brake. Equation 3.3 describes the calculation of brake force. $F_{\text{brakeTrans}}$ is the maximum applied force in N, which can be set in the configuration file. The brakeValue represents the engagement of the brake in the range from 0 to 1.

$$F_{\text{brake}} = F_{\text{brakeTrans}} * brakeValue \tag{3.3}$$

Equation 3.4 illustrates the calculation of the driving force. $F_{\text{gearTrans}}$ is a transmission factor that can be set in the configuration file. The gearValue is the physical state of the gear shift, and RPM are the revolutions per minute, respectively in $\frac{1}{60s}$. Furthermore, the degree of slope from the virtual environment and a configurable slope transmission factor in seconds are multiplied.

$$F_{\text{drive}} = F_{\text{gearTrans}} * gearValue * RPM * slope * slopeTransFactor \tag{3.4}$$

Equation 3.5 illustrates the calculation of the driving resistance force (F_{driveRes}). F_{air} describes the force of air resistance, F_{roll} rolling resistance, F_{slope} slope resistance and F_{acc} acceleration resistance each in N.

$$F_{\text{driveRes}} = F_{\text{air}} + F_{\text{roll}} + F_{\text{slope}} + F_{\text{acc}} \tag{3.5}$$

Equation 3.6 illustrates the combined force calculation. Variable drivEf is a configurable factor for drive efficiency, to model friction loss for instance.

$$F_{\text{total}} = \frac{F_{\text{drive}} - F_{\text{driveRes}} - F_{\text{brake}}}{drivEf} \tag{3.6}$$

The acceleration a is calculated by Equation 3.7. The physical mass m is the combined mass of bicycle and driver with the unit grams.

$$a = \frac{F_{\text{total}}}{m} \tag{3.7}$$

The current speed v is calculated with Equation 3.8 which contains the time t and the previous velocity v_0.

$$v = a * t + v_0 \tag{3.8}$$

3.5.3. Emotion Sensors

This PhD thesis focuses on the analysis of facial expressions and physiological data. The system for facial expression analysis was changed to ensure the required soft real time availabilities for a more reactive affective system (Requirement 2.5). Moreover, the system for physiological data acquisition was changed for this case study due the experiences of the first case study, described in Chapter 4.

The wearable body sensing platform biosignalsplux[7] was used. It provides EDA, blood volume pulse (BVP), ECG, piezoelectric respiration, and temperature sensors. The system operates at a sampling rate of 256Hz and all sensors are connected to the Plux hub. This hub transmits the data via Bluetooth to the server and thus ensures reasonable wireless data acquisition (Requirement 4.2.3).

In addition, the emotion sensor component was enhanced by a Tobii[8] X120 eye-tracker for preliminary experiments, which shows that Requirement 6.1.2 of an expandable system was met and . Further information about the analysis of the eye tracking data can be found in the master thesis of a team member (Bielenberg 2016). In addition, a thermal camera was mounted in front of the bike to measure thermal changes in the participant's face. First ideas of the data analysis are described in the report of Kletz (2016). The deployment of a Raytrix[9] R5 light field camera was evaluated, but did not prove helpful in the setup, further information can be found in Kletz and Kleimann (2016).

[7] http://biosignalsplux.com/
[8] https://www.tobiipro.com/
[9] https://www.raytrix.de/

3.5.4. Emotion Provoking Game Design

The emotion provoking and scene-based game of the second case study was refined to meet requirement 1.1 of the new game design, which enables multiple forms of adaptation (Requirement 1.1.9). The tailored game scenes were designed to steer participants in controlled emotional states. The linear structure was improved by a dynamic composition of game scenes (Requirement 1.1.11). A tunnel acts as a starting and ending portal to provide the transition between the scenes. It also enables dynamic game play and maintains the game flow for the user. The virtual tunnel is presented in Figure 3.13.

In the new game design the difficulty is configurable (Requirement 1.1.10), but was fixed for the second case study to limit the degrees of freedom in the experimental design. Although it may have been interesting to evaluate different variations, it had to be fixed to focus on the aims of this PhD thesis.

Figure 3.13.: Tunnel Design: Transition between Scenes

3.6. Discussion of the Experimental Setup

This chapter has demonstrated that the EmotionBike setup is able to use various sensors for emotion recognition, which is the ground work for the research questions presented in Section 3.3. The experimental results of two conducted case studies can be found in Section 4.5 and Section 5.3.

A physical exercise machine was used for an exergame. It was enhanced by steering abilities, a gear shift and a brake. The physical pedal resistance was adapted to the virtual environment as well as to the states of the gear and the brake. Moreover, the user's RPM while exercising was transmitted to the system to calculate the speed of the virtual bicycle.

In the following chapters the exergaming setup is applied to provide a proof of concept for the defined questions. This requires experiments with a focus on the different characteristics of the research questions. First a user study is conducted to evaluate if the system is able to provoke specific emotions by tailored game events in an exergaming context. This case study confirmed that facial expressions and physiological data can detect emotions in the exergaming context. Moreover, the influence of personality traits on the emotional reactions is analysed. The details are presented in Chapter 4.

A second case study is conducted that is based on the refined system design presented in this chapter. Thus new thrilling game elements are tailored, based on the new controls (gear and brake) and the emotion recognition is improved by real-time abilities. The details are presented in Chapter 5.

The system setup presented in this chapter combined with the case studies described in the next chapters are a proof of concept for the main research questions and will show that the system is able to provoke and detect emotions in an exergaming context. Furthermore, it will be shown that a dynamic story path based on users' emotions and actions can be applied.

4. Scheme and Experiment A: Novel Event-based Method for Facial Expressions and Physiological Data Analysis in Exergames

The field of entertainment computing continually creates new content to provide rich and unique experiences and engage players. The substantial progress in technology and storytelling facilitates new applications in various industry sectors, including healthcare, learning systems, cockpit designs and video games (Christy and Kuncheva 2014).

In learning systems for instance, it is important to detect the optimal challenge level to facilitate fast learning progress without exhaustion or frustration (Grafsgaard et al. 2013). Furthermore it has been shown that it is important to study frustration in affective games (Gilleade and Dix 2004), while cockpit scenarios focus on the recognition of mental overloads. Affective games alter their gameplay in response to the current emotional state of a player, and often provide biofeedback as described in Section 2.3, although a game can also be affective without offering biofeedback (Gilleade, Dix, and Allanson 2005). However, the first experiment focused on the provocation and measurement of specific emotions with the tailored game scenes; in the field of entertainment computing a scene-based game design is often used, as emotion recognition facilitate scene design on a personalised level (Togelius and Yannakakis 2016). In this work tailored game scenes feature game elements as stimuli to trigger specific emotions.

One challenge of designing entertainment systems is that the perception of entertaining values and the emotional reaction to similar game elements is highly individual. This results in a different personal experience for users. Thus, the presented system is designed to analyse these individual reactions and experiences. Furthermore, the system allows to conduct behavioural analysis studies. In the presented work a case study with eleven participants was conducted. The personal information in related research often includes a categorisation in personality traits to divide the participants in smaller groups, in order to decrease individual differences in the analysis. The widely used model of personality traits, the Big Five Model (John and Srivastava 1999) was applied in this work to decrease individual differences in the analysis. A detailed German version of a Big Five measurement, which was designed by Satow (2012) was utilised due to the links between physiology and personality described in the Chapter 2.5.1.

Biddle, Fox, and Boutcher (2000) linked physical activity to mood and emotions, and Lathia et al. (2016) to happiness. It has also been shown that interactive games are able to enhance the exercise experience (Warburton et al. 2007). Moreover, Hoda, Alattas, and El Saddik (2013) found specifically for cycling exergames that combining games with a physical bike increases the speed and the average RPM of participants by motivation, resulting in better exercise.

The analysis of emotional responses can aid game development and its design process by comparing alternative designs or studying the resonance and engagement of players. This PhD research developed two novel event-based emotion analysis methods. The first method is based on emotional reactions in facial expressions. The second enhances the facial expression analysis by physiological data. Moreover, two publications have shown that the system is able to provoke specific emotions with tailored game events, and that the system is able to measure emotions by analysing facial expressions and physiological data (Müller et al. 2015) and (Müller et al. 2016).

In summary, this chapter presents a new physical cycling game interface, that provokes and analyses emotions. A case study with eleven participants was designed and conducted to demonstrate that a provocation of specific emotions is possible with the tailored game events. The aim of the experiment was to prove the possibility of provoking specific emotions with the experimental setup, as well as to correctly measure the emotional state of a participant. The experimental setup was designed to examine individual differences in emotional reactions by categorising personality traits.

4.1. Experimental Procedure

The physical system setup for the experiment was described in Chapter 3. At the beginning of the experiment the participants were informed about the experimental procedure, as shown in Figure 4.1 and their right to abort the experiment at any time. They also had to sign a consent form. After that they were asked to fill out a couple of questionnaires with personal information, fitness level and game experience. A part of the personal information questionnaire included questions designed by Satow (2012), which allow for the categorisation according to the Big Five personality traits (openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism) as described in 2.5.1. The division of the participants into smaller groups was expected to decrease the individual differences in the analysis. Preliminary studies in our project team indicated that people with high values of extraversion displayed more emotional reactions.

Figure 4.1.: Experimental Process

The questions about their personal fitness level were aimed at their cycling experiences. A professional cyclist would probably display other physical reactions to the designed effort than someone who never rides a bicycle.

The sensors were placed on the participants as described in Section 3.2.3. To start the physical experiment, the participants had to mount the cycling exercise machine, subsequently the game was presented on a display in front of the bike. After the completion of all tasks the participants were asked about their physical strain perception on the range from 1 to 7, as defined by Borg (2004).

As described in Section 2.5, ethics approval is essential for affective computing studies. The experimental procedure presented in this thesis embraces the ethical guidelines of the UWS and was therefore authorised by the University's ethics committee.

4.1.1. Experimental Task

The researcher asked the participants to mount the exercise trainer and ride through six crafted games scenes on the virtual bicycle. In every scene the participants had to ride from the start to the finish line through the crafted game environment. The experimenter guided the participants with predefined phrases throughout the game to explain critical parts and objectives. Figure 4.2 illustrates the order of crafted game scenes described in Section 4.2. After every scene the participants were asked, which emotion they perceived.

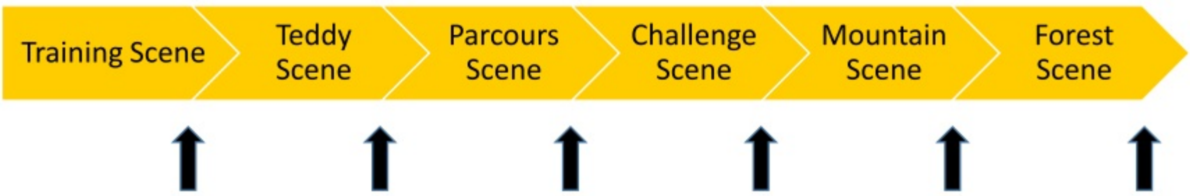


Figure 4.2.: Order of Game Scenes

4.1.2. Emotion Assessment

The emotion annotation is still an ongoing challenge in affective computing (Yannakakis, Martinez, and Garbarino 2016). As described in Section 2.5.1, it is not yet certain whether a self-assessment is sufficient or if an observer-assessment is preferable. Therefore, the experimental design includes the participants' self-assessment and the assessment of an independent human observer. The need for the observer to be trained has not yet been determined, hence the observer was trained, but no certificate was required. A categorised labelling of emotions was applied, as that has been described as more intuitive and is as acceptable as a ranking. The participants were asked to self-assess their emotions after each scene, as shown in Figure 4.3. The observer assessed the emotions directly in the scenes. An overall assessment at the end of the complete experiment would have presented some disadvantages, for instance the long time-span during which the participants would have to remember many different emotions. The provoked emotions might have influenced each other; a stressful element is, for instance, not perceived as that stressful in retrospect if was followed by an even more stressful element.

The presence of an independent observer may be perceived as a judgement and therefore be seen as a social-evaluative threat to maximise the stress provocation, as described in Section 2.2.2.

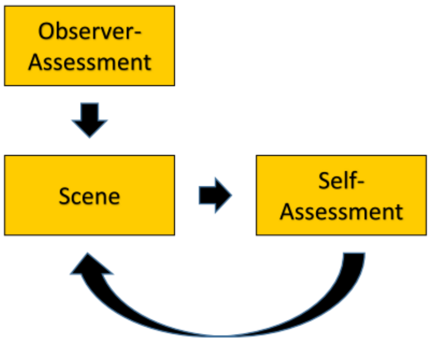


Figure 4.3.: Emotion Assessment Procedure

4.1.3. Methodological Explanations

The Living Place Hamburg is a smart home laboratory with a focus on future living concepts. This environment is very interesting for researchers, as well as for companies with an interest in future technologies. The author's supervising professor often invites people from science or business into the lab. The author of this thesis as well as her colleagues guided many tours through the lab for students or other guests. Moreover, many open house days took place and during these various events, the author was able to compile a list of volunteers. This list also includes the author's and other team members friends and sport colleagues. The participants were chosen randomly from this list, but were only recruited if they were available at the time planned for the experiments. In the first experiment most of the participants were colleagues from other offices or the sports club, due to the tight time schedule. The only condition was that the participant did not know the system from preliminary studies or exhibitions.

The average time a participant spent in the laboratory was approximately 60 minutes. This includes the sensor placement, the gaming tasks and the emotion assessment. However, it should be mentioned that the EEG sensor placement took around 15 minutes, due to the intricate technology. Every sensor needs to have contact, which is tested by a software. As previously mentioned the data were very noisy, and are therefore not further described in this thesis.

For ethical reasons, the participants were informed about the procedure before the experiment started.

4.2. Emotion Provoking Game Design

The designed game is a cycling game, which can be categorised as a fun-racer. It has no ambition to be physically accurate as a real-world bicycle. Instead it offers a broad range of gameplay mechanics and game events, which are not appropriate in realistic physical simulations. The game is designed according to the following requirements: intuitive (Requirement 1.1.6), easy to learn (Requirement 1.1.7), engaging (Requirement 1.1.4), and highly dynamic (Requirement 1.1.8). The game is controlled by the player, who has to physically accelerate and steer through the designed virtual environment. The virtual bike is simulated in soft real-time and the input is directly transferred into the game, hence the exercise machine is used as a game controller. Game elements are tailored to the needs of the experiments, and crafted scenes provide users with different exciting objectives (Requirement 1.1.5). In addition, tailored game scenes feature game elements as stimuli to trigger specific emotions (Requirement 1.1.1), for instance a crafted **Jump Scare Event** to provoke fear or surprise, which are logged into a database (Requirement 1.1.3). The start of the game scenes is software controllable (Requirement 1.1.2).

The game was designed to provide a moderate strain. The aim was an emotional provocation while exercising, but the game neither provides high intensity training nor sustained endurance exercise. The mean respiration rate was analysed as an indicator for moderate strain. Changes in skin temperature were also analysed in case they displayed interesting results.

The first experiment started with a **Training Scene** to become familiar with the controls.

This was followed by the **Teddy Scene**, the **Parcours Scene**, the **Challenge Scene**, the **Mountain Scene** and the **Forest Scene**, as illustrated in Figure 4.2.

Training Scene

In the **Training Scene** the virtual bicycle is placed on a street of a small town. The only task for the participants in this scene is to become familiar with the interface mechanics and the game world by driving from the start to the finish line. As described in Requirement 4.1.2, it is important to avoid emotional influences on the controls, and the training is designed to give the participant a chance to become familiar with them. Furthermore Requirement 1.1.7 is addressed, because it is easy to learn and does not require more than one training scene. The duration of the **Training Scene** was approximately 45 seconds. Furthermore, all sensor data were collected to ensure baseline possibilities.

Teddy Scene

The **Teddy Scene** is located in the same small-town environment as the **Training Scene**, although the street is populated with roaming teddy bears, as shown in Figure 4.4. These teddy bears can be avoided or hit, although in preliminary studies with the project team, killing the bears was shown to trigger feelings of joy. In case the virtual bicycle hits a bear, the teddy explodes and the **Teddy Hit Event** is sent to the message broker. After hitting a bear, the bike stops by setting the speed to zero and the player has to increase the speed again by pedalling fast. The duration of the **Teddy Scene** was approximately 50 seconds.

The **Teddy Scene** is designed to provoke joy and thereby to meet Requirement 1.1.1 of provoking a specific emotion.

Figure 4.4.: Roaming Teddy Bears on the Street in the **Teddy Scene** (Müller et al. 2015)

Parcours Scene

In the **Parcours Scene** presented in Figure 4.5 the participants have to collect 20 coins, which are arranged as a parcours traversal with traffic cones, shown in Figure 4.5a. If the participant does not collect every coin and cross the finish line, the bike is teleported back to the start line. After the collection of a coin the **Coin Collection Event** is triggered and a particle effect of smileys appears, as shown in Figure 4.5b, which was designed to provoke surprise. The **Coin Collection Event** was designed to provoke joy, since it is a step to fulfil the task. The scene does not end before all coins are collected and the finish line is crossed. The duration of **Parcours Scene** was approximately one minute and 30 seconds. The **Parcours Scene** is designed to provoke joy or surprise and thereby to meet Requirement 1.1.1 of provoking specific emotions.

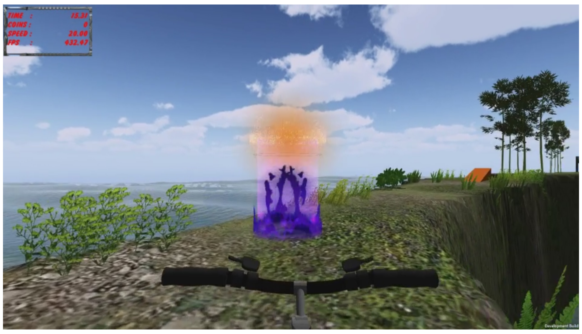
(a) Coin Collection in the **Parcours Scene**

(b) Smileys after Coin Collection

Figure 4.5.: Pictures of the **Parcours Scene** (Müller et al. 2015)

Challenge

The **Challenge Scene** starts on a small path on a mountain road. There is a cliff on the left and on the right side of the road, as shown in Figure 4.6. The player needs to jump over a giant gap, shown in Figure 4.6b, in order to finish the scene. To fulfil this task a booster gate, shown in 4.6a, has to be crossed in proper alignment with a ramp to reach the necessary speed to catapult the bike over the gap. A successful landing on the other side, as shown in Figure 4.6c, is a challenging task, due to the alignment and steady steering. Thus the ramp's width is increased with the number of trials. To provoke frustration, every time the player falls down the cliff the **Falling Event** is triggered and the scene starts again. It is also possible that the user will express joy; many people smile in frustrating situation, as described in Section 2.4.3. The sensitive steering usually leads to a frustrating number of trials, therefore an EDA response is expected. The duration of **Challenge Scene** was approximately five minutes, depending on the number of trials. The **Challenge Scene** is designed to provoke frustration and thereby meets Requirement 1.1.1 of provoking a specific emotion. In addition the integrated task provides an exciting challenge to the participants (Requirement 1.1.5).



(a) Booster Gate and Ramp

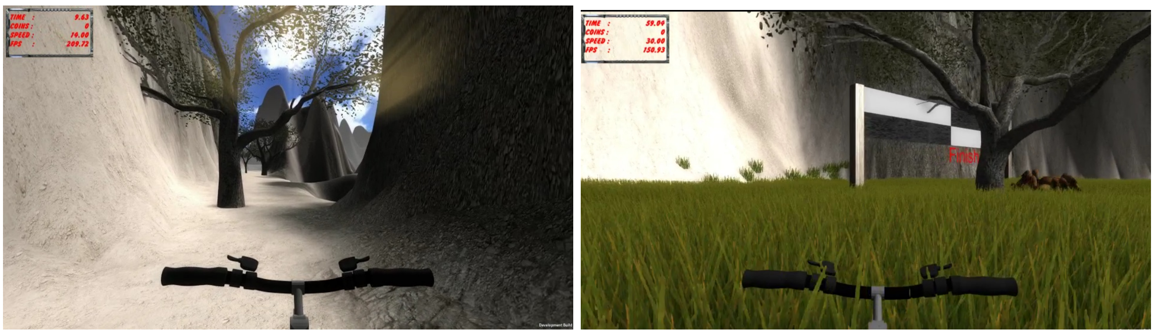
(b) Jump over the Giant Gap

(c) Landing on the Plateau

Figure 4.6.: Pictures of the **Challenge Scene**, (a) (Müller et al. 2015), (b) (Müller et al. (2016), ©2016 IEEE)

Mountain Scene

The **Mountain Scene**, shown in Figure 4.7, starts on a mountain road, presented in Figure 4.7a. The participants have to climb a hill with the bike. The physical resistance of the ergometer pedals increases in relation to the degree of ascent and thereby builds upon Requirement 3.1 of providing a controlled pedal resistance. The influence of physical strain is expected to increase the respiration rate of the participants, as Plarre et al. (2011) found respiration features are related to physiological stress. After reaching a plateau, shown in Figure 4.7b, which is filled with tall and dense grass, the bike is attacked by spiders. The **Spider Attack Event's** are triggered to surprise the player (Requirement 1.1.1), as the spiders are hard to spot due to the dense environment. In addition, wall traps emerge unexpectedly and block the bike for a short period of time. The duration of **Mountain Scene** was approximately one minute and ten seconds.



(a) Mountain Road (b) Plateau Filled with Dense and Tall Grass

Figure 4.7.: Pictures of the **Mountain Scene**, (a) (Müller et al. (2016), ©2016 IEEE)

Forest Scene

In the **Forest Scene** the virtual bike is located in a dark forest, which the participants have to cross. The only light source is the bicycle headlight and self-illuminated coins, which guide the player through the forest. Before the player enters the finish line a **Jump Scare Event** is triggered to provoke surprise or fear (Requirement 1.1.1). All bike controls are immediately disabled and monsters spawn in front of the bike. This event additionally includes a horrible shout. The monsters are shrouded in red lights and a grainy film effect distorts the vision, as shown in Figure 4.8. The **Jump Scare Event** is expected to provoke a response in EDA data. The duration of **Forest Scene** was approximately 45 seconds.

Figure 4.8.: **Jump Scare Event** in the **Forest Scene**, (Müller et al. (2016), ©2016 IEEE)

Table 4.1 provides an overview of emotional provoking game scenes and their objectives.

Scene	Event	Objective	Target Emotion	Target Reaction
Teddy	Teddy Hit	Avoid or hit roaming teddy bears	Joy	
Parcours	Coin Collected	Collect all coins	Joy	
Challenge	Falling	Make it to the finish line	Joy, Frustration	EDA response
Mountain	Spider Attack	Ride up a hill	Surprise	Increased respiration rate, EDA response
Forest	Jump Scare	Cross a dark forest	Surprise, Fear	EDA response

Table 4.1.: Overview of Emotional Provocation in Game Scenes

4.3. Participants' Profile

Eleven participants took part in the case study. Three were female and eight male, and their age ranged from 19 to 41, with an average age of 27. Seven of them described themselves as casual gamers and said to play video games two to three hours a week. Four declared not to play video games on a regular basis. Seven participants confirmed to have experience with modern game controls like Wii remote or Kinect.

All participants stated that they engage in sports, seven of them on a regular basis. Nine participants stated they cycle on a weekly basis, ranging from one hour to four hours. None of the participants said they do professional cycling exercise training, participate in spinning classes, or have a cycling exercise machine at home.

4.4. Data Analysis Method

The timestamps for the emotion provoking game events were logged in a database and the sensor data near to the tailored events were considered for the analysis. In this work facial expressions and physiological data were analysed. According to Gunes et al. (2011b) there is no unique answer for an optimal window size for data analysis in related literature; moreover, the provided window size differs between modalities. Thus an appropriate analysis window was defined for each sensor in this thesis.

The evaluation of affective systems is usually based on self- or observer-assessments (Calvo and D'Mello 2010), thus this work applies a combination of both. The facial expressions of the participants were assessed by observers during the experiment.

The observers were trained by works of Ekman (2004), but they did not have a professional certificate. In the first experiment a student co-worker was trained for this task.

Subsequent to completion of each level the participants were asked to self-assess their emotions in a retrospective manner. In order to address the phenomenon mentioned by (Grafsgaard et al. 2013) among others, namely that facial expressions vary between individuals, the results were compared with the self and observer-assessments.

4.4.1. Novel Analysis Method for Facial Expressions in Exergames

Paul Ekman developed a facial action coding system (FACS), which allows to train people to recognise emotions in faces, as described in Section 2.4.1. FACS can be used to label images used as training input for computer systems based on machine learning. The facial expression data in this work are analysed with the Computer Expression Recognition Toolbox (CERT) introduced by Littlewort et al. (2011), which is based on Ekman's theory. CERT provides a probability for basic emotions: joy, disgust, anger, fear, neutral, sad, surprise and contempt. It is based on machine learning algorithms to train the recognition rate in image sequences showing FACS coded emotions, as described in Section 2.4.1. The calculation of probability values was described for the predecessor of CERT, Emotient and is described in Equation 5.1.

Section 3.1 states that the game should provoke different emotions. The analysis of emotions therefore focuses on the evaluation of the aimed emotions, which should be triggered in a specific scene. Furthermore the self-assessment and the observer-assessment were analysed and in case they differed from the provocation aim, the described emotion was analysed in more detail as well.

The probability values near the provoking events were compared with the target emotions within a time window of 3 seconds (0.5 seconds before and 2.5 seconds after an event). Those data were evaluated for emotion probability peaks. An example event analysis is shown in Figure 4.9. For each target event maximum and mean basic emotion probability were calculated. Provoked emotion responses were taken into account for the summary results if they exceed a threshold of 0.5 or more specifically 50 % (Müller et al. 2015).

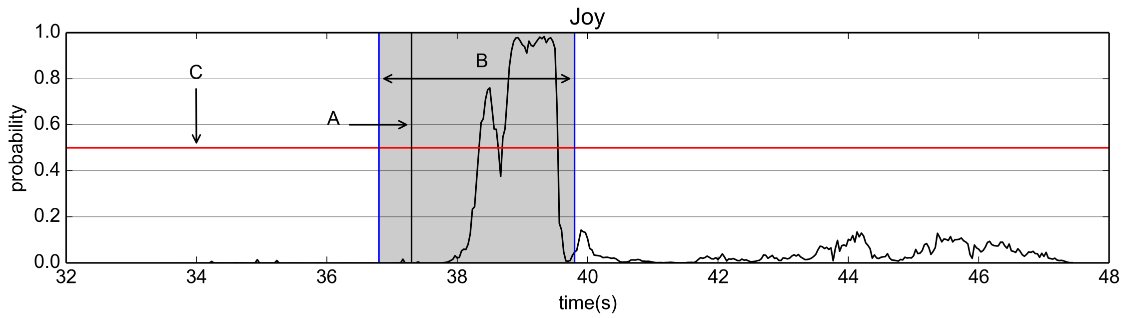


Figure 4.9.: Example Emotion Analysis for Data of **Forest Scene**: Black vertical line marks a game event (A), interval (B) shows the analysis window, red line (C) marks the detection threshold, black line shows an emotion response curve. (Müller et al. 2015)

4.4.2. Novel Analysis Method for Physiological Data in Exergames

Various applications benefit from research with stress recognition (Wallbott and Scherer 1991), for instance driving (Healey and Picard 2005) or pre-hospital care (Kindness, Mellish, and Masthoff 2013b). It has been shown that mental strain may increase negative emotions (Feldman et al. 1999), lead to discriminative facial expressions (Wallbott and Scherer 1991) and self-reports (Baggett, Saab, and Carver 1996). Therefore, this section illustrates the extension of the facial expression analysis method by physiological data analysis. Emotion classification methods cannot be directly applied to detect stress, as they focus on a specific set of emotions and do not include the whole variety of negative emotions that are able to provoke stress (Plarre et al. 2011).

The presented combined analysis is designed to evaluate physical and cognitive stress as well as emotional reactions. In this way the integration of physiological data analysis allows to enhance the entertaining exergame system by means of a balanced workout programme.

The physiological system reacts to more than stress, as described in Section 2.4.2. Changes in the physical or mental condition are also influential, hence the recognition of physiological responses during physical activity differs from recognition without movement. The designed exergame system requires physical effort, therefore the presented approach focuses on an event-based analysis, similar to the facial expression analysis method presented in Section 4.4.1.

In related literature, the analysis of EDA, respiration and temperature data has shown to be very promising, as described in Section 2.4.2. EDA is a part of the autonomous nervous system, which is known to be closely associated with the arousal of the participant (Boucsein 2012). The EDA signal was used for basic emotion recognition (Jang et al. 2011). In the presented PhD thesis the recognition of basic emotions is achieved through the analysis of facial expressions. The findings of the developed facial expression analysis method have shown individual differences in the expression of emotions. Many expressed emotions neither correlated with the self-assessments nor with the observer-assessments. The results of the facial expression analysis will be described in more detail in Section 4.5.3. The integration of physiological data analysis promises to enhance the successful emotion recognition rates. The recognition of frustrating game elements is expected to benefit from a combination of EDA data and facial expressions. Furthermore, the analysis of EDA data can be applied for stress detection, as described in Section 2.4.2.

Stress can be measured by baseline differences in the EDA data (Wallbott and Scherer 1991), but the results are controversial (Friedrichs et al. 2015). The developed analysis method is event-based, but for the sake of completeness baseline differences have been evaluated as well. Thus, in the presented analysis method, EDA sensor data were evaluated in two ways. An activity peak detection for the sensor data focused on the phasic component of the signal and a baseline evaluation focused on the tonic component of the signal. Furthermore the event-based analysis method was enhanced by a peak detection algorithm. Related research has shown that the activity response in the skin conductance signal can be delayed by one to three seconds after an event. In addition, the ascent of a curve peak can take up to five seconds (Boucsein 2012). Therefore this work uses an expansive analysis window. To choose an applicable window size, two events were analysed in detail. Table 4.2 presents the peaks detected for the different analysis windows regarding the **Falling Event** in the **Challenge Scene**. It shows that an analysis window from one second before to ten seconds after the occurrence of the event is preferable for this scene. Table 4.3 shows differences in recognised peaks for different analysis window sizes for the **Jump Scare Event** in the **Forest Scene**. In this event, an analysis window from one second before to eight seconds after the occurrence of the event is enough. Further results are presented in Section 4.5.4.

Table 4.2.: Correspondences between the *Falling Event* in the *Challenge Scene* and EDA peaks detected (Total Matches:70), (Müller et al. (2016), ©2016 IEEE)

Left Border (s)	Right Border (s)	Window Size (s)	Matches Found	Matched (%)
1.0	4.0	5.0	60	86
1.0	6.0	7.0	64	91
1.0	8.0	9.0	66	94
1.0	10.0	11.0	68	97
2.0	4.0	6.0	62	89
2.0	6.0	8.0	66	94
2.0	8.0	10.0	68	97
2.0	10.0	12.0	70	100

Table 4.3.: Correspondences between the *Jump Scare Event* in the *Night Scene* and EDA peaks detected (Total Matches:11), (Müller et al. (2016), ©2016 IEEE)

Left Border (s)	Right Border (s)	Window Size (s)	Matches Found	Matched (%)
1.0	4.0	5.0	4	36
1.0	6.0	7.0	8	73
1.0	8.0	9.0	10	91
1.0	10.0	11.0	10	91
2.0	4.0	6.0	4	36
2.0	6.0	8.0	8	73
2.0	8.0	10.0	10	91
2.0	10.0	12.0	10	91

4.5. Experimental Results

4.5.1. Personality Traits

Before the experiment started, the participants were asked to fill out a questionnaire designed by Satow (2012). This questionnaire allows a categorisation according to the Big Five, the personality traits extraversion, agreeableness, conscientiousness, neuroticism and openness to experiences. The results of the analysis are presented in Figure 4.10. Three have a high intensity of agreeableness and four have a high intensity of neuroticism. Two participants showed a low intensity of agreeableness, seven a low intensity of conscientiousness and four a low intensity of neuroticism. Two participants showed a low intensity of openness to experiences while two can be categorised as very open to experiences.

Three participants are highly extroverted, while two are introverted. An interesting finding is that the highest peaks in EDA data and the facial expression analysis was exhibited by extroverted participants. One of the participants who were classified as extroverted had the highest value in the facial expression analysis and the highest peak in the EDA analysis. Another highly extroverted participant exhibited high values in facial expressions, and the third participant had very high peaks in EDA analysis. However, this effect requires more participants to obtain significant results. Moreover, no negative correlation between EDA and neuroticism, as has been reported in literature (Brouwer et al. 2013), was found in the case study.

4.5.2. Emotion Assessment

The emotion assessment of the first case study consisted of self-assessment and the assessment of an independent observer. Both are illustrated in Appendix 4.5.2. The self-assessments of the **Training Scene**, the **Teddy Scene** and the **Parcours Scene** are presented in Table 4.4. The results of the **Challenge Scene**, the **Mountain Scene** and the **Forest Scene** are shown in Table 4.5. The observer-assessment is presented in Table 4.6 and Table 4.7, respectively.

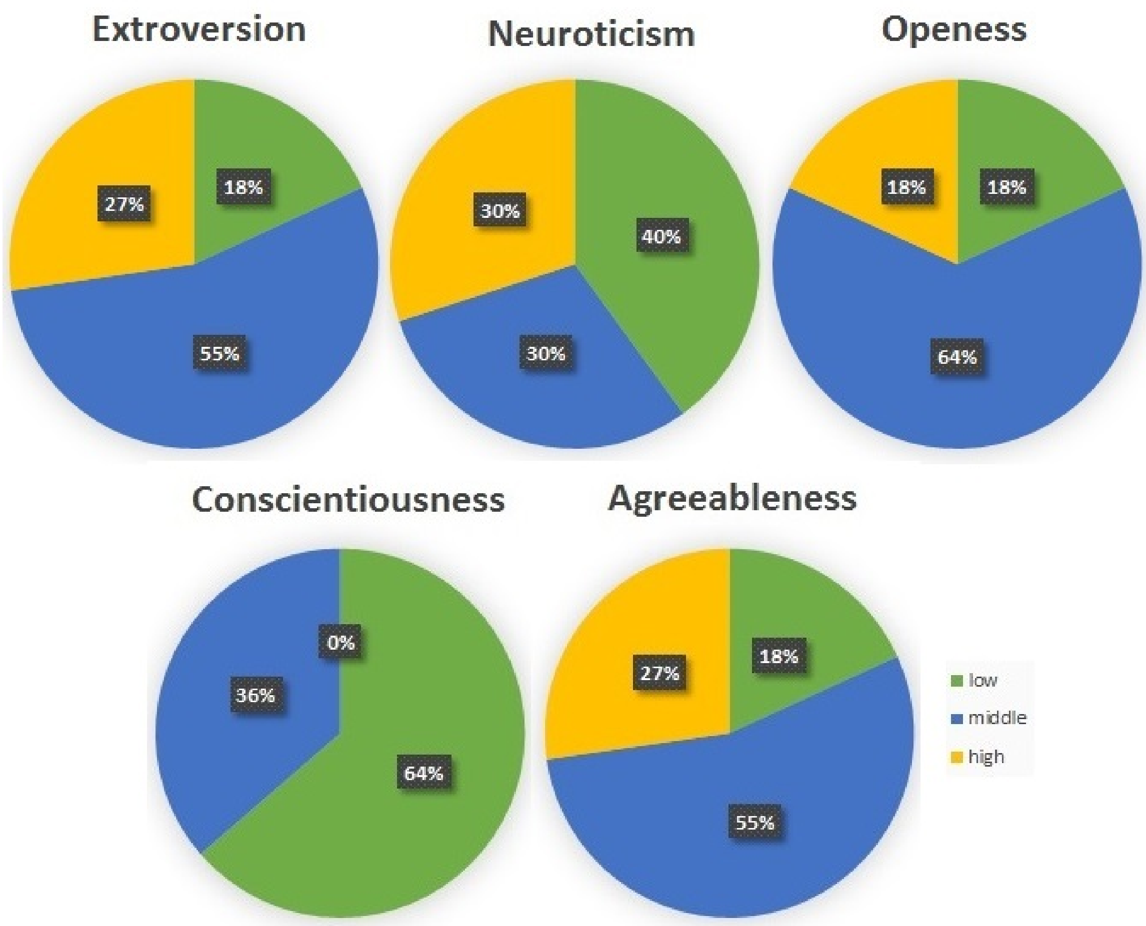


Figure 4.10.: Proportional Distribution of the Big Five Personality Traits of Eleven Participants

Participant Id	**Training**	**Teddy**	**Parcours**
1	Neutral	Surprise, Confused, Joy	Ambition, Joy
2	Neutral	Relaxed/Calm	Bored
3	Joy	Compassion, Joy	Stressed out
4	Concentrated/Interested Expectant	Joy	Concentrated/Interested Expectant
5	Neutral	Joy	Confused
6	Neutral	Joy Concentrated/Interested	Neutral
7	Concentrated/Interested	Joy	Neutral
8	Joy	Joy	Joy, Ambition
9	Neutral	Joy, Neutral	Concentrated/Interested
10	Joy	Joy	Joy, Effort/ Strain
11	Concentrated/Interested Joy	Joy Relaxed/Calm	Frustrated Joy

Table 4.4.: Participants' Self-Assessment for **Training Scene**, **Teddy Scene** and **Parcours Scene** during the First Case Study

Participant Id	Challenge	Mountain	Forest
1	Stressed out, Relief	Effort/ Strain, Surprise	Discomfort/ Anxiety/ Tension, Frightened/ Shocked
2	Stressed out, Frustrated	Effort/ Strain, Joy	Frightened/ Shocked
3	Stressed out	Effort/ Strain, Surprise	Frightened/ Shocked
4	Stressed out, Ambition	Confused	Frightened/ Shocked
5	Angry	Relaxed/Calm	Stressed out
6	Surprise, Ambition	Effort/ Strain	Frightened/ Shocked
7	Frustrated, Stressed out, Ambition	Effort/ Strain	Concentrated/Interested, Frightened/ Shocked, Fear
8	Joy, Discomfort/ Anxiety/ Tension, Ambition	Joy, Effort/ Strain	Frightened/ Shocked
9	Joy, Frustrated, Ambition	Bored, Effort/ Strain	Frightened/ Shocked, Joy
10	Ambition, Effort/ Strain	Surprise, Effort/ Strain, Frightened/ Shocked	Frightened/ Shocked
11	Ambition	Confused	Neutral

Table 4.5.: Participants' Self-Assessment for **Challenge Scene**, **Mountain Scene** and **Forest Scene** during the First Case Study

Participant Id	**Training**	**Teddy**	**Parcours**
1	Joy	Confused, Joy	Concentrated/ Interested
2	Joy, Confused	Neutral	Concentrated/ Interested, Bored
3	Neutral	Joy, Effort/ Strain	Stressed out
4	Concentrated/ Interested	Joy	Concentrated/ Interested
5	Concentrated/ Interested	Joy	Concentrated/ Interested
6	Relaxed/Calm	Surprise	Relaxed/Calm
7	Joy, Concentrated/ Interested	Joy, Frightened/ Shocked	Concentrated/ Interested, Effort/ Strain
8	Joy, Concentrated/ Interested	Joy	Joy, Concentrated/ Interested
9	Joy	Joy, Concentrated/ Interested	Concentrated/ Interested, Confused
10	Joy, Surprise, Concentrated/ Interested	Joy, Concentrated/ Interested, Ambition, Expectant	Concentrated/ Interested
11	Concentrated/ Interested	Joy	Joy, Ambition

Table 4.6.: Observer-Assessment for **Training Scene**, **Teddy Scene** and **Parcours Scene** during the First Case Study

Participant Id	Challenge	Mountain	Forest
1	Concentrated/ Interested, Surprise, Joy, Stressed out	Effort/ Strain, Joy	Surprise, Frightened/ Shocked
2	Surprise, Joy, Stressed out	Effort/ Strain, Surprise	Concentrated/ Interested, Stressed out, Frightened/ Shocked
3	Frustrated, Effort/ Strain	Effort/ Strain, Joy	Frightened/ Shocked
4	Confused	Bored, Confused	Frightened/ Shocked
5	Surprise, Stressed out	Neutral	Neutral
6	Frustrated, Ambition	Relaxed/Calm	Surprise
7	Joy, Concentrated/ Interested, Relief	Joy, Surprise, Discomfort/ Anxiety/ Tension	Frightened/ Shocked, Neutral
8	Joy, Stressed out, Effort/ Strain	Joy, Effort/ Strain, Frightened/ Shocked	Concentrated/ Interested, Frightened/ Shocked
9	Joy, Concentrated/ Interested, Frustrated, Relief	Bored, Stressed out, Effort/ Strain	Joy, Concentrated/ Interested, Frightened/ Shocked
10	Joy, Surprise, Concentrated/ Interested, Frustrated, Effort/ Strain	Joy, Surprise, Disgust, Effort/ Strain	Surprise, Concentrated/ Interested, Frightened/ Shocked
11	Joy, Frustrated, Ambition	Concentrated/ Interested	Surprise

Table 4.7.: Observer-Assessment for **Challenge Scene**, **Mountain Scene** and **Forest Scene** during the First Case Study

Training Scene

The **Training Scene** was designed to allow the participants to become familiar with the controls, thus no events were triggered during the scene. Five of the participants stated they felt neutral and four that they felt joy. The observer recognised joy six times. In addition the observer saw concentration by six participants, which were exactly the anticipated results.

Teddy Scene

In the **Teddy Scene** ten self-assessments indicated joy. These results were amplified by nine observer-assessments, stating joy. These results were as expected, as it was designed to provoke joy.

Parcours Scene

In the **Parcours Scene** eight observer-assessments indicated concentration or interest, although only two participants stated they felt concentrated or interested. Four self-assessments described joy during the scene. None of the assessments showed surprise, which was expected from the smiley particle effect. The positioning of the coins and the traffic cones may have provoked concentration instead of the intended joy.

Challenge Scene

The repetition of **Falling Events** during the **Challenge Scene** was described as frustrating by three of the participants, five of them stated they felt stressed out and seven mentioned ambition. Four observer-assessments indicated stressed out and five frustration. Ambition was recognised twice by the observer.

Mountain Scene

Eight self-assessments of the **Mountain Scene** and six observer-assessments indicated effort or strain. These results were expected, as the participants had to cycle up a hill at the beginning of the scene. Only three of the participants stated they were surprised by the **Spider Attack Event** and only three observer-assessments indicated this.

Forest Scene

The **Jump Scare Event** in the **Forest Scene** was described as shocking or frightening by nine of the participants. Eight of the observer-assessment indicated frightened or shocked; in addition surprise was recognised twice.

Summary of Emotion Assessment

Table 4.8 summarises the results of the self-assessment and the observer-assessment. It provides a comparison of the assessment related to the game scenes.

Scene	Emotion	Occurrence of Self-assessment	Occurrence of Observer-assessment
Teddy	Joy	10	9
Parcours	Joy	4	2
Parcours	Surprise	0	0
Parcours	Concentrated / Interested	2	8
Challenge	frustration	3	5
Challenge	Stressed Out	5	4
Challenge	Ambition	7	2
Mountain	Effort / Strain	8	6
Mountain	Surprise	3	3
Forest	Shocking / Frightening	9	8

Table 4.8.: Summary of Emotion Assessment

4.5.3. Facial Expression Analysis

The facial expressions occurring during the **Training Scene** were not analysed, as no events were triggered to provoke emotions. The aim of this scene was to become familiar with the controls of the virtual bicycle. Therefore the results of the other scenes are presented in the following. Furthermore, the results of the personality traits regarding the facial expression analysis are presented.

Personality Traits

Participants with a high value in extraversion had a higher maximum value in the expression of joy. The maximum probability for extroverted person was 0.89 against 0.76 for introverted participants.

Teddy Scene

In the **Teddy Scene** the **Teddy Hit Event** was designed and crafted to provoke joy. The results of the facial expression analysis for this event showed that this worked for eight out of ten participants (one participant avoided to hit a bear). Figure 4.11 shows the highest probability for the emotion joy, as max value, which was expressed by a participant during all events. In this scene joy should be provoked and therefore only the results for joy are displayed. The mean value is calculated over all events and in this scene many teddy bears can be hit. One of the participants (Participant Id 8) did not hit any teddy, due to successful avoidance. Participant number two did not recognise the **Teddy Hit Event**, as the teddy bumped into the virtual bicycle from behind, resulting in low values. The relative high max values of participant four in combination with a low mean value indicate that he or she felt joy at least once.

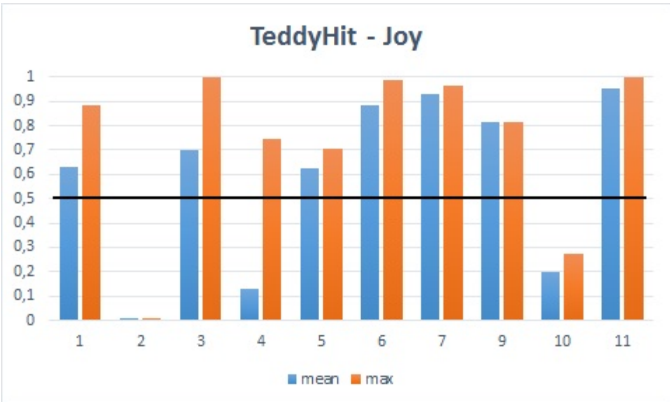
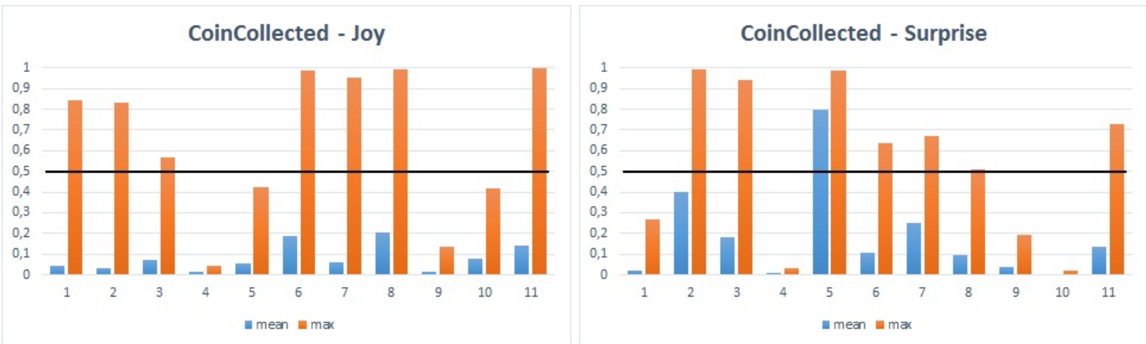


Figure 4.11.: Experimental Results of the Facial Expression Analysis of the **Teddy Hit Event** in the **Teddy Scene** (Müller et al. 2015)

Parcours Scene

In the **Parcours Scene** the participants faced the task to collect 20 coins in order to complete the scene successfully. The **Coin Collection Event** was expected to create joy, as every collected coin is a step to fulfil the task. Furthermore, the **Coin Collection Event** triggered a smiley particle effect, which might provoke surprise. Figure 4.12a show that seven of the participants displayed joy in their facial expressions at least once. The low mean value indicates that the provocation effect was not maintained for the duration of 20 coins. In addition the player was teleported to the start line in case he or she did not collect all the coins, which also influenced the emotional reactions. Near misses in collection and anticipation of finishing the scene provoked a stronger response.

The results of the facial expression analysis for surprise is presented in Figure 4.12b. Six participants displayed high surprise values during the **Coin Collection Event**, indicating that they were surprised at least once. In this scene joy or surprise should be provoked and therefore only the results for joy and surprise are displayed.

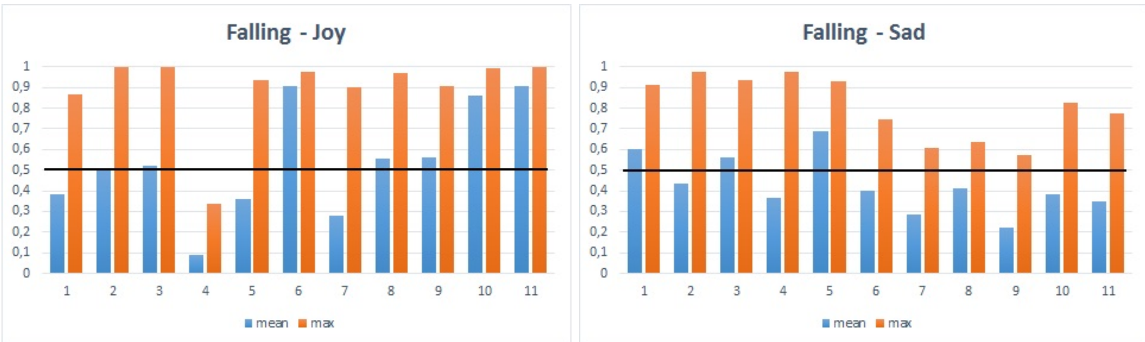


(a) Experimental Results of Joy Probabilities (b) Experimental Results of Surprise Probabilities

Figure 4.12.: Experimental Results of the Facial Expression Analysis of the **Coin Collection Event** in the **Parcours Scene** (Müller et al. 2015)

Challenge Scene

In the **Challenge Scene** the participants had to cross a booster gate to jump over a giant gap, which is a very challenging task. Many attempts are necessary, as landing on the other side requires correct steering alignment. Every time the player falls down the cliff, the **Falling Event** is triggered. The continual repetition is intended to provoke frustration. The ambition of most participants was very high, resulting in various emotional reactions. Ten out of eleven participants displayed joy during the **Falling Event** and all of them displayed sadness at least once, as shown in Figure 4.13a and Figure 4.15d, respectively. The phenomena that people smile in natural frustrating situations has been reported by Hoque and Picard (2011). In this scene frustration should be provoked, but CERT does not provide probability values for frustration. For this event therefore joy and sadness have been evaluated. Sadness as a reference for a negative emotion and joy, as many people smile during natural frustration.



(a) Experimental Results of Joy Probabilities (b) Experimental Results of Sad Probabilities

Figure 4.13.: Experimental Results of the Facial Expression Analysis of the **Falling Event** in the **Challenge Scene** (Müller et al. 2015)

Mountain Scene

In the **Mountain Scene** the participants encountered spiders roaming in the tall and dense grass. The **Spider Attack Event** was not perceived as very surprising. Only four participants displayed surprise in their facial expression, as shown in Figure 4.14. No mean values are presented in the diagram, as the **Spider Attack Event** is not repeated. Many of the participants seemed distracted, due to the high physical resistance while cycling up the mountain at the beginning of the scene. In this scene surprise should be provoked and therefore only the results for surprise are displayed.

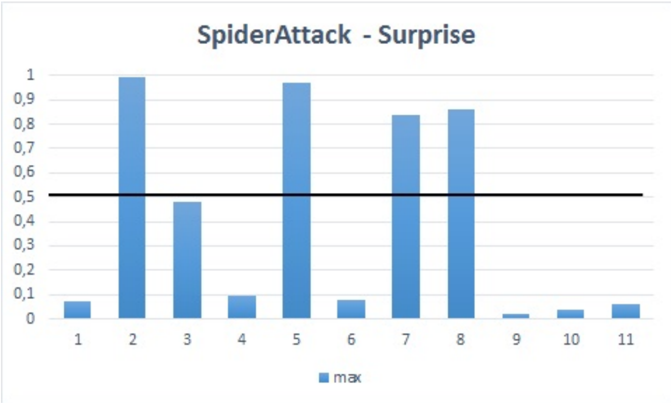


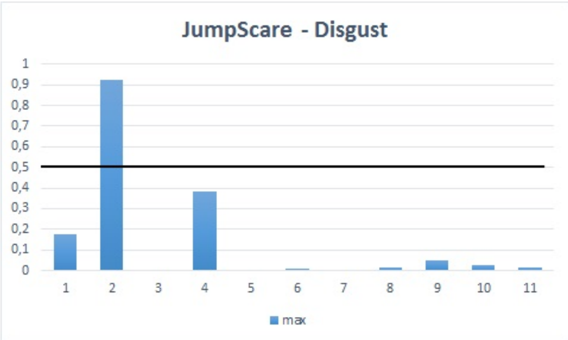
Figure 4.14.: Experimental Results of the Facial Expression Analysis of the **Spider Attack Event** in the **Mountain Scene** (Müller et al. 2015)

Forest Scene

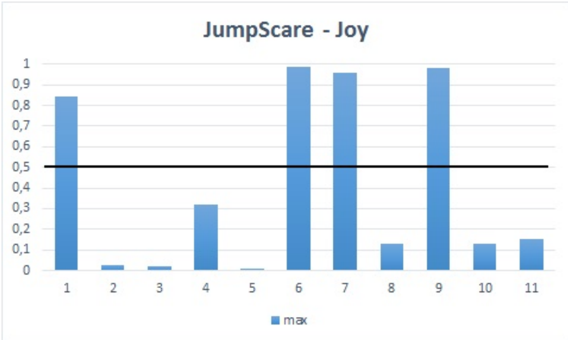
The **Jump Scare Event** in the **Forest Scene** provoked the most varying emotional responses in the facial expressions, as shown in Figure 4.15. Two of the participants showed fear and one displayed surprise. Four showed joy and one participant had high values for disgust. These results neither correlate with the self-assessment nor with the observer-assessment as described in Chapter 4.5.2. As the **Jump Scare Event** occurred only once, no mean values are presented. In this scene surprise or fear should be provoked, but no clear results were found. High responses were found in fear and surprise, but also in disgust and joy.

(a) Experimental Results of Fear Probabilities

(b) Experimental Results of Surprise Probabilities



(c) Experimental Results of Disgust Probabilities



(d) Experimental Results of Joy Probabilities

Figure 4.15.: Experimental Results of the Facial Expression Analysis of the **Jump Scare Event** in the **Forest Scene** (Müller et al. 2015)

Conclusion of Facial Expression Analysis

The experimental results of the presented novel facial analysis method demonstrated that the tailored game events are able to provoke specific emotions displayed in the participants' facial expressions. Table 4.9 provides an overview of the results of provoked emotions per scene. The crafted **Teddy Hit Event** and the **Falling Event** provided the best results in provoking specific emotions. The **Coin Collected Event** and the **Jump Scare Event** provoked a more varied emotional response in the type and probability of emotion.

Scene	Event	Target Emotion	Participants Provoked in %
Teddy	Teddy Hit	Joy	80
Parcours	Coin Collected	Joy	63.6
Challenge	Falling	Joy	90.9
Challenge	Falling	Sad	100
Mountain	Spider Attack	Surprise	36.4
Forest	Jump Scare	Surprise	18.2
Forest	Jump Scare	Joy	36.4
Forest	Jump Scare	Fear	18.2
Forest	Jump Scare	Disgust	9.1

Table 4.9.: Summary of Provoked Emotions with crafted Events for the Eleven Participants (Müller et al. 2015)

4.5.4. Physiological Data Analysis

Respiration Rate

The experiment was designed to provide moderate strain for the participants, as the aim was an emotional provocation while exercising, and not high intensity training or sustained endurance exercise. The respiration rate can increase up to 50 breaths per minute (Nakajima, Tamura, and Miike 1996) during exercises in which a high effort is required. The results presented in Figure 4.16 show that the aim of provoking moderate physical strain was successful. The mean respiration rate did not increase much over the scenes (Rounded: 14.6 to 15.3). The maximum respiration rate was shown in the **Mountain Scene** with a value of 22.

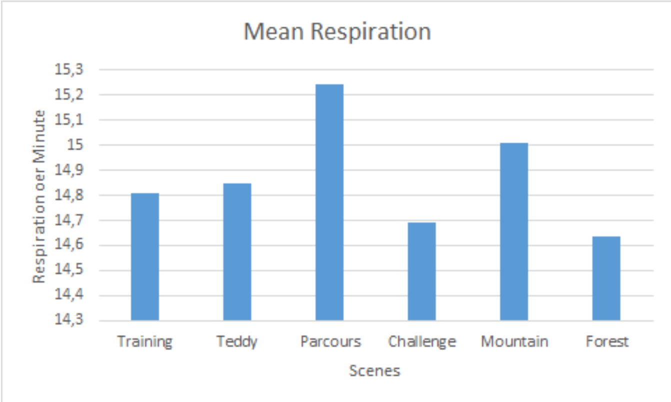


Figure 4.16.: Experimental Results of the Physiological Data Analysis: Mean Respiration Rate for all Scenes

EDA Baseline

A slow drift of the baseline over time has been reported by Kächele et al. (2015). The tonic component of the EDA signal is not highly diagnostic in the conducted case study due to the required physical activity. Figure 4.17 shows that the baseline increases from scene to scene. For most subjects the absolute values changed due to the physical effort. However, the presented event-based approach is robust to these changes, as it focuses on the phasic component of the signal and therefore on a peak based evaluation.

Training Scene

The **Training Scene** was implemented without emotion provoking events, as the only aim was to become familiar with the controls and the virtual game environment. The results of the physiological data analysis show a mean respiration rate in the **Training Scene** of 14.8. Eight of the participants had a high mean value, which might be explained by nervousness or the challenge of becoming familiar with the controls.

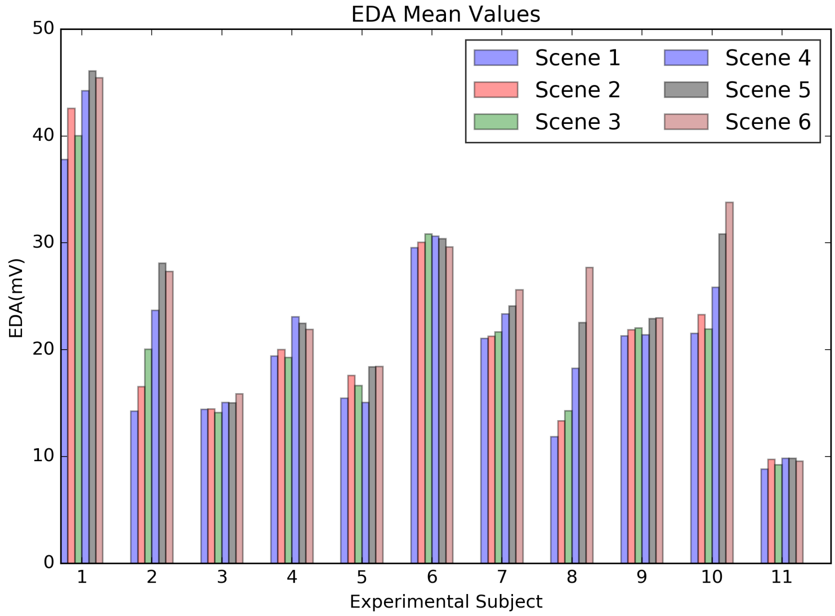


Figure 4.17.: Experimental Results of the Physiological Data Analysis: EDA baseline, (Müller et al. (2016), ©2016 IEEE)

Teddy Scene

Joy was displayed in the facial expressions of most participants. The EDA data analysis did not show results, which confirms the hypothesis that the smile in the **Challenging Scene** differs from the smile in the **Teddy Scene**, due to the lack of frustration provocation. The mean respiration rate in the **Teddy Scene** was 14.9.

Parcours Scene

For the **Coin Collection Event** in the **Parcours Scene** no interesting results was found in the EDA data. However, an interesting finding was that the mean respiration rate in the **Parcours Scene** was the highest for all scenes with a value of 15.3.

Challenge Scene

In the **Challenge Scene** the participants needed to cross a booster gate to reach the finish line, which triggered the **Booster Event**. Due to the challenging steering alignment most of the attempts lead to falling off the cliff, which triggers the **Falling Event**. An interesting finding was that seven participants had positive peaks in their temperature change data by crossing the booster gate for the first time and four had negative peaks. However, this is not a significant result and further evaluations of this effect would be required. Examples of temperature change data are presented in Figure 4.18.

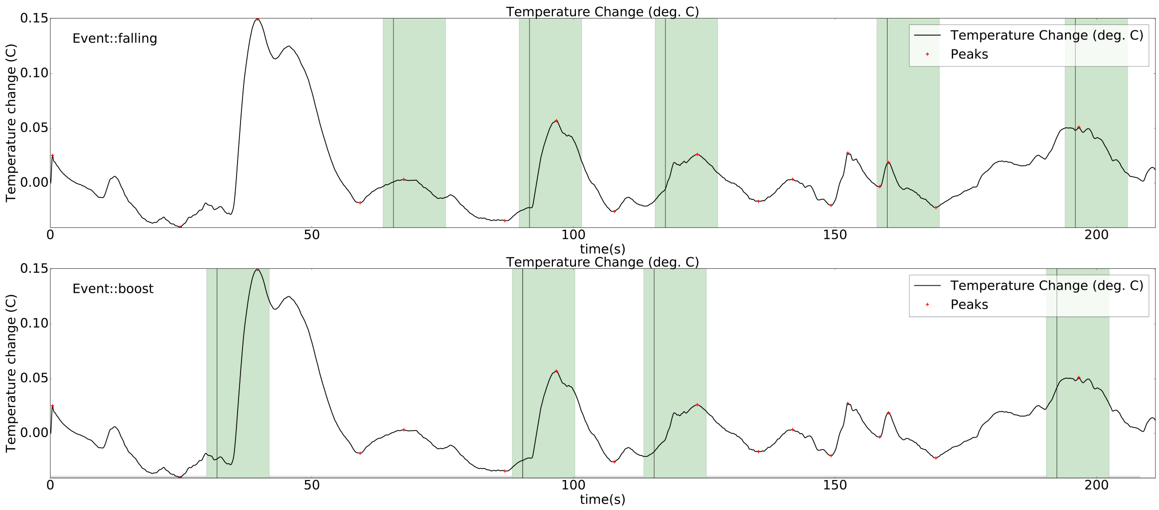


Figure 4.18.: Temperature change raw data during the **Challenge Scene** for the **Boost Event** and the **Falling Event**, (Müller et al. (2016), ©2016 IEEE)

A matching performance of 97% in EDA data was found for an analysis window of one second before and ten seconds after the event. This effect can be interpreted as tension due to the challenging objective. As previously described, the facial expression analysis was not very successful in this scene due to individual reactions. As shown in Figure 4.19, a combination of the event-based EDA data analysis with the event-based facial expression analysis is very promising, as high tension was provoked in the presented case study due to the repetition of falling off the cliff, which starts the scene again. The mean respiration rate in the **Challenge Scene** was 14.7 breaths per minute, which is related to moderate strain.

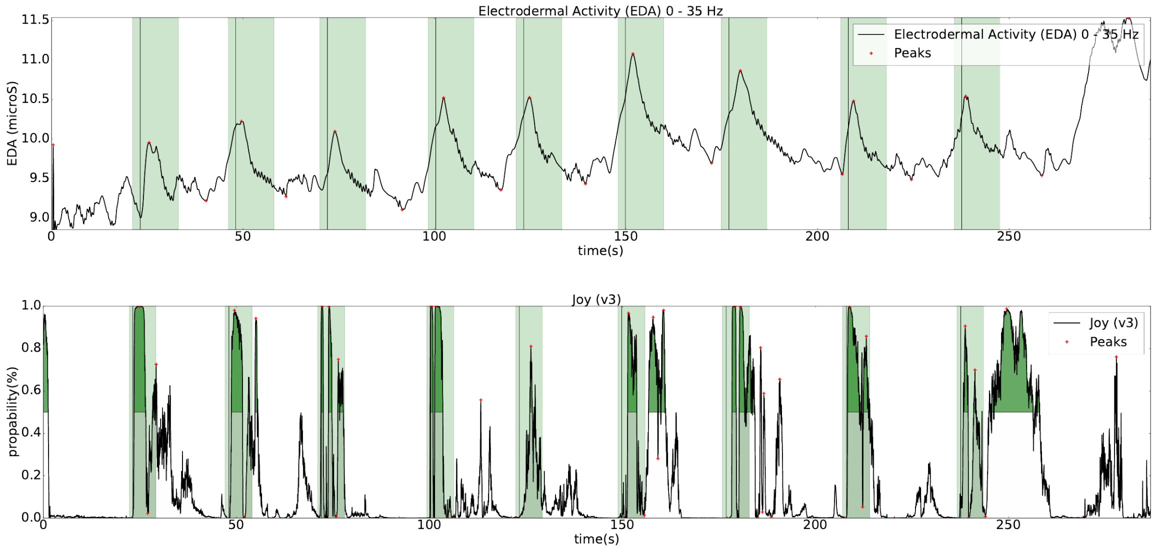


Figure 4.19.: EDA raw data and Joy Output during the **Challenge Scene** for the **Boost Event** and the **Falling Event**, (Müller et al. (2016), ©2016 IEEE)

Mountain Scene

The basic emotions theory does not include physical strain, which the CERT software can thus not recognise in the facial expressions. The physiological data analysis shows that the mean respiration rate increased to a value of 15.0 in the **Mountain Scene**, due to the increasing physical strain while climbing the hill.

Forest Scene

The physiological data analysis results were much more promising than the facial expression analysis, in which varying emotional responses were found. Ten out of eleven participants showed a high peak in the analysis window of the **Jump Scare Event**, resulting in a recognition rate of 91%. Example EDA raw data are presented in Figure 4.20. Furthermore, seven participants had a positive peak in the data of the temperature change sensor and two exhibited a negative peak. The mean respiration rate in the **Forest Scene** was 14.6.

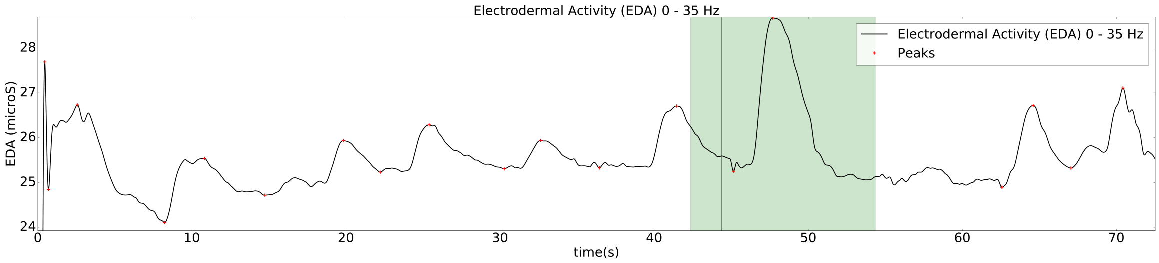


Figure 4.20.: EDA raw data during **Forest Scene** for the **Jump Scare Event**, (Müller et al. (2016), ©2016 IEEE)

Conclusion Physiological Data Analysis

The physiological data analysis enhances the previously described facial expression analysis method. EDA responses were found for 97% of the **Falling Event** and for 91% of the **Jump Scare Events**. The integration of physiological data analysis in the **Challenge Scene** helps for instance to avoid false positives in joy recognition. The analysis of EDA data is preferable for the presented exergaming setup, as the consideration of the respiration rate and the temperature change are less expressive. Table 4.10 summarises the results of physiological reactions.

Scene	Event	Target Reaction	Results
Challenge	Falling	EDA response	97%
Mountain	Spider Attack	Increased respiration rate	Highest max value in respiration rate (22) no considerable EDA responses
Forest	Jump Scare	EDA response	91%

Table 4.10.: Overview of Physiological Provocation Results

4.6. Conclusion

In this chapter, the benefits of a computerised exercise machine to provoke specific emotions in game contexts have been established. The presented affective exergame was designed to provoke specific emotions by means of tailored game events. An experiment was designed and a case study with eleven participants was conducted to prove that the system is able to provoke and measure user emotions.

The experimental procedure included the evaluation of personality traits to obtain more insights into individual differences in emotional reactions, which can vary between groups of users. Therefore emotion responses may be correlated with factors such as personality traits, gender or age to enable an improved prediction of emotional reactions, which was indicated by the results of the first case study. The highest peaks in EDA data and facial expressions were exhibited by extroverted participants for the **Falling Event** and the **Jump Scare Event**. One of the participants classified as extroverted displayed the highest values in facial expressions and EDA data. To get more significant results, the number of participants will be increased in the next experiment.

A novel event-based analysis method for facial expressions was developed, and the results have shown (Müller et al. 2015) that the tailored game elements are able to provoke the predefined emotions. The **Teddy Hit Event** and the **Falling Event** provoked emotions in most subjects. The **Coin Collection Event** and **Jump Scare Event** provoked a more varied response in type and probability of emotion. The aim of this experiment was more than to provoke the specific emotion, but rather to show that emotions can be detected with the presented setup.

A novel analysis method for physiological data was presented to enhance emotion recognition. It was shown in Müller et al. (2016) that the combination of physiological data analysis with facial expression analysis magnifies the potential for affective exergaming applications. Furthermore the identification of individual differences in the perception of entertaining values increased by adding physiological data analysis. The emotion detection rate for the **Jump Scare Event** in the **Forest Scene** for instance could be enhanced by physiological data analysis. Moreover, the detection rate of the emotional responses to **Falling Event** in the **Challenge Scene** could be increased to 97%.

Game development is one application area that benefits from this research. The provided analysis of game events can aid the design process by comparing alternative designs or studying the resonance and engagement of players. Furthermore, the presented analysis methods provide a quantified measurement of emotional user reactions to analyse user experience. A further interesting application for the presented research is healthcare systems, but more flexibility is required to provide a balanced workout programme, which will be the focus of the next experiment.

A more reactive affective system that is able react to users' emotions has to analyse the emotions in soft real-time. The applied software for facial expression recognition, CERT, does not have these abilities. Thus, for further experiments another tool needs to be purchased and integrated to setup a more adaptive affective system. The tools available on the market were evaluated by another team member Bernin et al. (2017).

Sometimes it was difficult for the participants to describe their emotions. For the next experiment a list of suggestions was defined, which was visible to the participants during the assessment. For instance, in the German language there is the word for surprise (Überraschung), which was used for a positive surprise (Erstaunt) and for scary frightening events (Erschrocken). Thus the list of suggestions included both terms, and the participants were asked which one they had in mind. Moreover, there are many emotions with a similar meaning in the German language, thus for every scene there was a list of perceived emotions, although they often described the same emotion. A list of suggestions will overcome the difficulties of emotion categorisation in the post processing of the experiment as well.

Some of the participants had trouble with the physiological sensors, as the cable was tangled up in the ergometer pedals. Therefore the next experiments required wireless data transmission of the sensors.

The experimental setup was part of a classical laboratory setting (Figure 3.4), which often makes people feel uncomfortable. Further experiments require a more natural environment, thus for the next experiment a smart home laboratory, the Living Place Hamburg was chosen, to provide more comfortable surroundings. The Living Place enhanced the situational context awareness of the system by the smart home environmental control (Brauer 2014; Ellenberg et al. 2011).

5. Scheme and Experiment B: Novel Real-time Multimodal Emotion Analysis Method to Support a Dynamic Story Path Based on Individual Emotional Reactions to Multi-emotion Provoking Game Scenes

The paradigm of "healthy living" is attracting growing worldwide attention. The increasing sedentary lifestyle in western countries requires a well-balanced workout programme combined with adequate nutrition (Hamilton et al. 2012). Exergames have been shown to support a healthier lifestyle, as entertaining content offers excellent possibilities to encourage endurance by motivating elements (Malaka 2014) to support a personalised fitness programme.

Therefore expectations increased for entertaining content in healthcare and fitness applications during the last years. The market benefits from recent technological enhancements in fitness assessment, such as quantified-self devices for activity tracking and physiological data acquisition. As previously mentioned, ergometers are often used in medical rehabilitation and for endurance or high intensity indoor fitness training. Thrilling and exciting content facilitates user motivation, thus the physical cycling exergame controller was enhanced by a gearshift and a brake to facilitate the design of more natural game elements by the provided controls. However, for healthcare and rehabilitation applications it is important to consider personal fitness and to avoid physical or mental stress. Thus, the presented entertaining content includes a mixture of physiological stress and emotion provoking game elements, which are controlled in intensities to provide an exciting but not taxing exercise experience. The designed exergaming system encourages an ambient personal fitness programme through entertaining content.

To automatically detect the optimal challenge level is as important in learning systems as in training systems, in order to provide a sustained activity or facilitate a fast learning progress without exhaustion (Malaka 2014) or frustration (Grafsgaard et al. 2013). Adaptive learning systems are designed to provide this optimal challenge by considering emotional engagement (Alyuz et al. 2016). Affective games that provide an appropriate challenge level can be highly motivating, but need to ensure to avoid a mental overload, similar to learning systems. In exergames, adaptation of the required effort can avoid a taxing effect.

Today's video games utilises open worlds, nonlinear game designs and action-based character designs to overcome linear story lines and static experiences. These advances aid the design of individual and evolving narratives. It is more entertaining to integrate individual emotional reactions in the content selection, which goes beyond a choreographed content. Moreover, the difficulty level of video games can be adapted to a person's game experience, resulting in improved motivation. Many commercial games (Christy and Kuncheva 2014) were developed during the last decade, demonstrating the physiological and behavioural effects of affective computing. For the second case study, the analysis method described in Chapter 4 was enhanced by a real-time emotion evaluation, to allow the system to adapt the gameplay dynamically.

Due to the virtual game environment and the physically accessible exercise machine controller, situational context awareness was ensured and enhanced by changing the location of the physical cycling exergame controller to the smart home environment, the Living Place Hamburg (Brauer 2014; Ellenberg et al. 2011).

Physical aspects combined with emotional states have a high potential to enhance user motivation and training effects. In this chapter an affective entertainment application is presented, which is able to take participants on an **Emotional Journey** by adapting the gameplay to individual emotional responses. The term **Emotional Journey** is applied for entertainment computing. It is only loosely related to other areas, such as therapy, motivational or design research, in which the term might also be used. The concept of an **Emotional Journey** is introduced to modify the often applied character-based game development to a gameplay that is affected by emotional responses supplemental to action-based modifiers. The **Emotional Journey** is designed as the emotional path a player takes during the entertainment. The journey includes established emotion provoking game scenes. They are combined with new entertaining content. The emotions are recognised by the developed emotion recognition analysis methods and the gameplay is adapted to ensure the provocation of the specific emotions in the right order.

The new game design provides a multi-emotion provoking and continuous game design, which makes this study an important step on the way toward holistic integrated exergames, and goes beyond the previously designed single-emotion provoking game scenes. This work shows the potential of individualised content to enhance users' endurance by means of increased motivation, and is a first step toward a personalised adaptive fitness system.

An experiment was designed and a case study with 25 participants was conducted to showcase the potential of the enhanced exercise experience by multi-emotion provoking game elements and an emotion responsive gameplay. The results were published in Müller et al. (2017a) for the **Emotional Journey** and in Müller et al. (2017b) for the multi-emotion provoking game elements.

5.1. Experimental Procedure

The experimental procedure was similar to the first case study, as shown in Figure 4.1, and embraced the same ethical guidelines. The experiment started with an explanation of the experiment, and the sensors were placed before the tasks began. The physiological sensors were extended by an ECG and a BVP sensor. The sensor positions were based on Einthoven's triangle, as described in (Zalis and Conover 1972). The anode was placed on the right clavicle, the ground electrode on the left clavicle, and the cathode on the left forearm. The BVP sensor was attached to the third finger of the left hand and the Plux connection hub was mounted at the belt of the participants. The experimenter who guided the sensor placement was of the same gender as the participants to avoid feelings of discomfort (Requirement 2.2).

The experimental procedure included the same questionnaire about Big Five personality traits as the first case study. However, the analysis was conducted by another team member.

Every scene started and ended with a tunnel. This was an appropriate place to ask the participants for a short self-assessment, to ensure a brief retrospection time for perceived emotions. The observer was asked to assess the participants' emotions during the scene and again had the task to function as a judge, as it has been shown by Dickerson and Kemeny (2004) that social-evaluative threats are perceived as stressful, as described in Section 2.2.2, to meet Requirement 1.2 and intensify the provocation.

5.1.1. Emotion Assessment

The emotion assessment was based on a short self-assessment after each scene and an observer-assessment during the ride, based on the results of the first case study. The self-assessment was enhanced, due to the findings from the first case study. A predefined list of emotions that were relevant for the experiment was presented to the participants. Figure 5.1 shows these German terms for emotions. As previously mentioned, it is quite challenging to find a suitable translation. Essentially the list included basic emotions, and those that might be perceived by the participants during the crafted game scenes.

Following a completed scene, the participants were asked about the fun they perceived, using the discrete scale presented in Figure 5.2. This scale was introduced by Nacke et al.

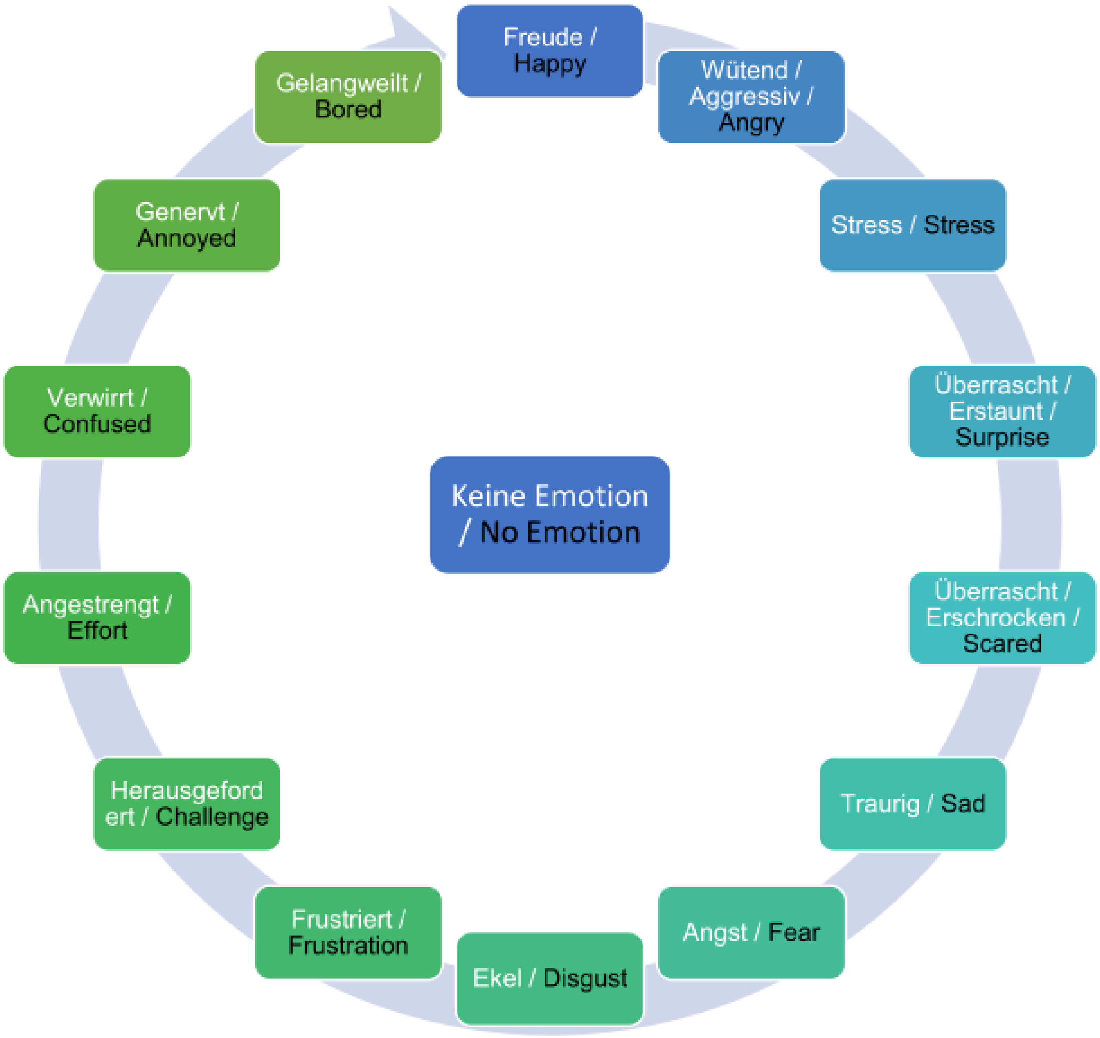


Figure 5.1.: Relevant Emotions for the Second Case Study

(2011), although it was translated to German for this experiment. The scale ranges from 1 (not much fun) to 5 (very fun).

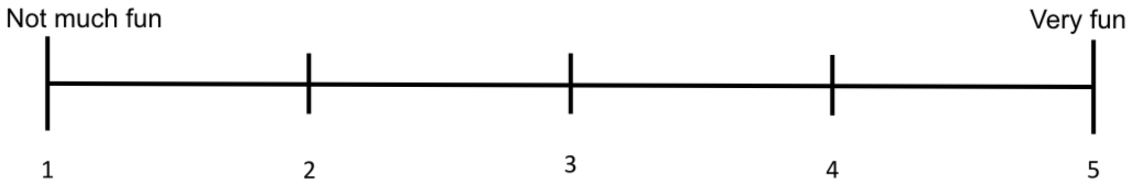


Figure 5.2.: Scale for Perceived Fun

5.1.2. Methodological Explanations

In this experiment the participants were chosen in a similar way to the procedure described in Section 4.1.3. The experiences of the first experiment suggested that the author should not conduct the experiment with volunteers that she knew personally to avoid any influences. Therefore many friends and flatmates of the other team members were invited. The final decision about the participation of the volunteers was based on the tight time schedule, and the required minimum number of 20 participants. In the end 25 participants were invited to participate on a voluntary basis. None of the participants in this experiment had taken part in the first experiment.

After the participants arrived in the smart home laboratory the environment was first explained to them, as in most cases they were very impressed and had many questions. Then they were informed about the procedure, so they did know that the experiment aims at emotion provocation. This embraces the ethical guidelines of the UWS and was therefore authorised by the University's ethics committee.

The time a participant spent in the laboratory was approximately one and a half hours. This includes the introduction, the questionnaires, the sensor placement, the time of exercising and the assessment between the scenes and after the three tasks of the experiment.

5.1.3. Experimental Tasks

The experiment was divided into three parts. The first part was the **Emotional Journey** to showcase the dynamic system control (Requirement 1.1.11). The second part consists of newly designed multi-emotion provoking game scenes. These advanced game scenes included different kinds of emotion provoking element to craft a more natural game experience. In the **Frozensea Scene** different kinds of emotion provoking elements were designed to evaluate the intensity. In addition, the scene provides a challenge to the participants. In the **Cliff Scene** different stressful game elements were combined to analyse the provocation intensity. It also provides a challenging task to the participants and is located in an inconvenient environment. New exciting game content was created by utilising the new possibilities of the controls (Requirement 1.1.13).

In the last part the participants had to climb a mountain with the virtual bicycle. The physical resistance of the ergometer pedals was set to a very high value of 150 Watt, for half of the group. The pedal resistance of the control group was set to a maximum of 100 Watt to evaluate the combination of physical stress and emotional provocation.

Experiment Part 1: Emotional Journey Concept

The **Emotional Journey** part of the experiment was designed to showcase the system's ability to adapt the game play to the participants' individual emotional responses in addition to their actions (Requirement 1.1.9).

Based on previous findings, three established scenes were chosen to serve as mandatory scenes in the **Emotional Journey**: the **Forest Scene**, the **Challenge Scene** and the **Teddy Scene**. The emotions that were expected to be provoked by the scenes are surprise or fear in the **Forest Scene**, ambition, stress or frustration in the **Challenge Scene** and joy in the **Teddy Scene**.

In case the **Forest Scene** was not able to provoke fear or surprise, the **Explosion Scene** was started. The **Downhill Scene** was started to provoke stress or frustration in case the **Challenge Scene** did not perform well. Moreover, the **Glade Scene** was designed to provoke joy if the **Teddy Scene** did not work out.

This experimental task evaluates if emotional responses flatten by multiple event occurrences. The **Jump Scare Event** in the **Forest Scene** was presented to the participants multiple times to analyse the hypothesis of such a flattening effect, which was indicated by preliminary studies.

At the beginning of the experiment and before the scenes for the **Emotional Journey** started, the emotional and physical conditions of the participants were assessed. The emotional condition was assessed with the SAM questionnaire, as described in Section 2.5.1. In addition the participants were asked about their current emotional state on a range from 1 (not good) - 9 (very good). The Borg scale (Borg 2004) was applied to assess the participants' physical condition, as it is a validated scale in sport sciences. Following the **Emotional Journey** the assessment procedure was repeated, as presented in Figure 5.3.

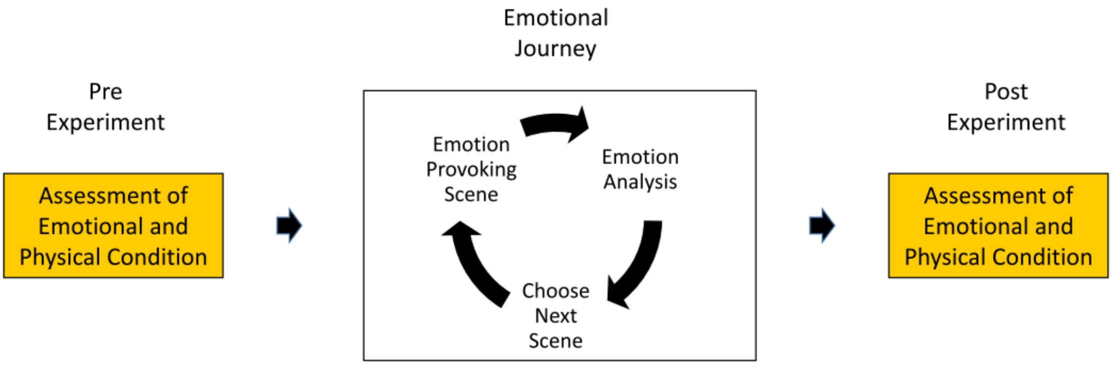


Figure 5.3.: Emotional Journey Process

The facial expressions and physiological data were evaluated during the scene and a decision was made after the participants reached the finish line in the transition portal tunnel. The experimenter in the experimental control station was the one in charge of the final decision about the success of a provocation, as he had a full overview about the emotional responses of all the applied emotion sensors in combination with the assessments.

An adaptation was made to increase the successful provocation rate, as illustrated in Figure 5.3. In addition the **Emotional Journey** concept was designed to showcase the ability of the system to adapt to individual emotional responses.

Established Scenes:

The new design of the **Forest Scene**, shown in Figure 5.4, supports a dynamic software controllable **Jump Scare Event**, shown in Figure 4.8, instead of a fixed place of occurrence (Requirement 1.1.12). In the previous experiment it was a singular event, but in the second case study the **Jump Scare Event** was triggered three times by the experimenter to measure the flattening effects of emotions. The controls were not disabled after the **Jump Scare Event** occurred, to exclude a fixed break.

Figure 5.4.: **Forest Scene**

For the **Challenge Scene**, shown in Figure 4.7, a more natural ramp, shown in Figure 5.5, was designed, and a magnetic effect was implemented to have direct control of the landing on the other side. This made it possible to control the player's success or failure by means of software (Requirement 1.1.12). Moreover, a software controllable function was designed to teleport a player to the finish line in case he or she could not complete the task.

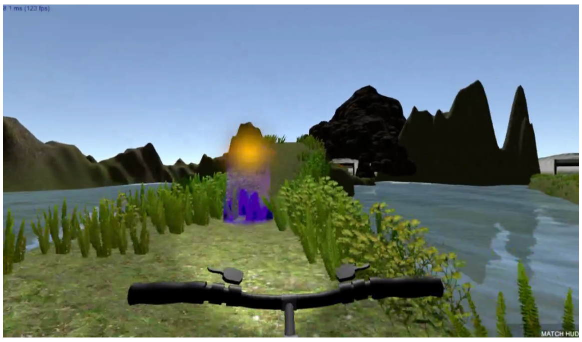


Figure 5.5.: More natural Ramp in the **Challenge Scene**

Figure 5.6 displays the scene map of the **Challenge Scene**.

In the **Teddy Scene**, the number of teddy bears is configurable in the new game design. In the second case study many more teddy bears were roaming the street, to ensure the

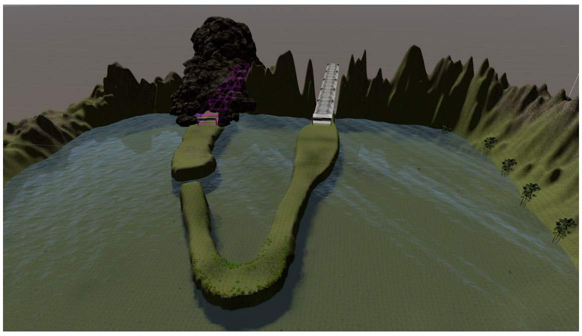


Figure 5.6.: Scene Map of the **Challenge Scene**

participants were not able to avoid all the bears, shown in Figure 5.7. In case the participant hit a teddy, it exploded in a bloody animation, and his body remained on the street.

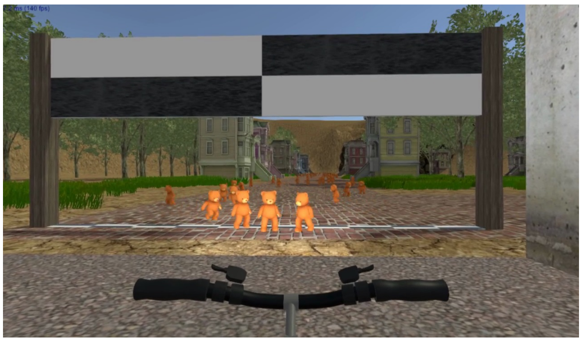


Figure 5.7.: More Teddy Bears in the **Teddy Scene**

Figure 5.8 displays the scene map of the **Teddy Scene**. No teddy bears are roaming the street in the scene map.

Additional Scenes:

The **Fork Scene**, shown in Figure 5.9 substitutes the **Training Scene** of the previous case study. The aim of this scene is to become familiar with the controls. Furthermore, the participants were asked by the researcher to test the gearshift and brake inside the scene. The participants could decide if they wanted to go left, with nice green trees, or right, which looks

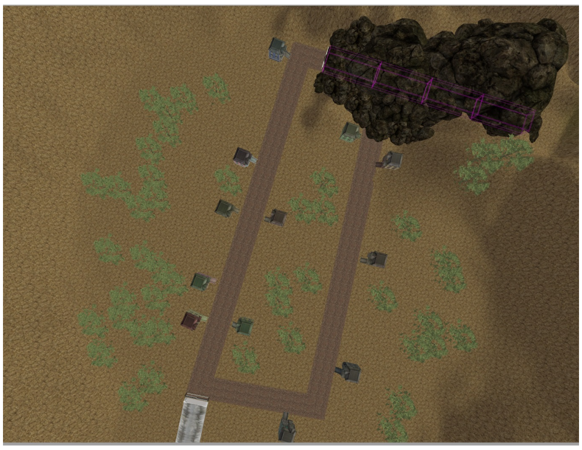


Figure 5.8.: Scene Map of the **Teddy Scene**

more like a wild desert of a hilly landscape. The duration of **Fork Scene** was approximately one minute.

Figure 5.9.: **Fork Scene**

Figure 5.10 displays the scene map of the **Fork Scene**.

The **Explosion Scene**, shown in Figure 5.11, starts in a desert with much military equipment. This scene suggests to be set on a battlefield. During the ride through the threatening environment the software controllable **Explosion Event** is triggered followed by a loud sound to provoke surprise or fear. The player dies and has to start the scene again in case the virtual bicycle is too close to the centre of the explosion. Which of the positioned explosion is triggered to control the **Player Died Event** is also software controllable (Requirement

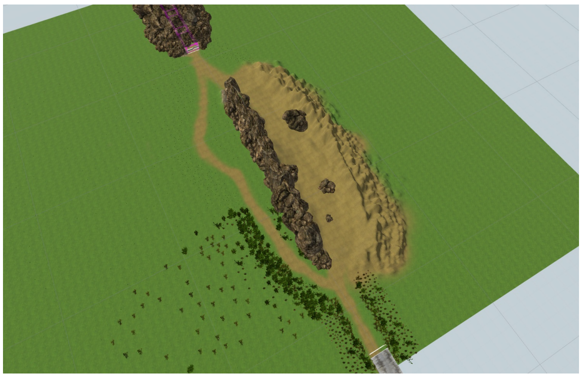


Figure 5.10.: Scene Map of the **Fork Scene**

1.1.12), but in the second case study all explosions were triggered by the experimenter, to maximise the scary effect. The duration of **Explosion Scene** was approximately one minute, depending on the number of player died events.

Figure 5.11.: **Explosion Scene** (Müller et al. (2017a), ©2017 IEEE)

Figure 5.12 displays the scene map of the **Explosion Scene**.

In the **Downhill Scene** the virtual bicycle is placed on top of a mountain. The only way down is a narrow ramp with a wood design, as presented in Figure 5.14. The bike speeds up due to the negative slope, and in front of the finish line there is a sharp curve without railings, shown in Figure 5.14. In case the player falls off the ramp into virtual lava the **Player Died Event** is triggered and the scene starts again. In this scene the brake is disabled to increase the difficulty and thereby meet Requirement 1.1.13 of applying the new game

Figure 5.12.: Scene Map of the **Explosion Scene**

controls. The scene is designed to be very challenging, due to the high speed, the sharp turn and the disabled brake. The duration of **Downhill Scene** was approximately three minutes, depending on the number of trials. Figure 5.13 displays the scene map of the **Downhill Scene**.

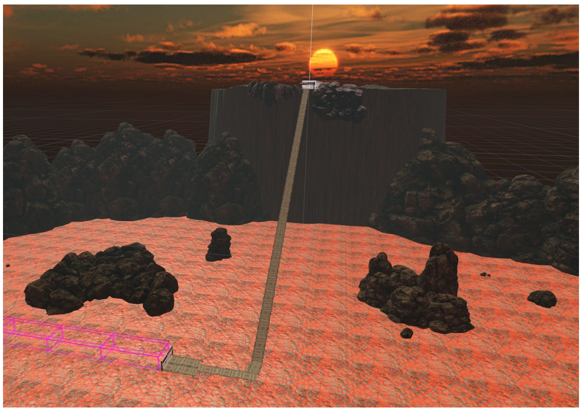
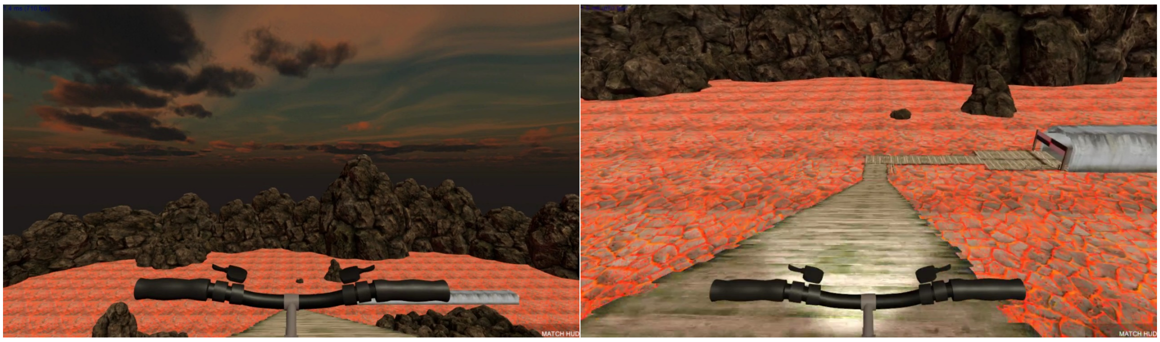


Figure 5.13.: Scene Map of the **Downhill Scene**

The **Glade Scene** is a very nice environment with beautiful plants and trees designed as a glade of a forest. It has a very calming design to provide a relaxing atmosphere. During the ride through the glade the participants encounter a funny big rabbit, shown in Figure 5.15. The duration of **Glade Scene** was approximately one minute and 30 seconds. Figure 5.16 displays the scene map of the **Glade Scene**.



(a) Virtual Lava in the **Downhill Scene** (b) Sharp Curve in the **Downhill Scene**

Figure 5.14.: **Downhill Scene** (Müller et al. (2017b), ©2017 IEEE)

Figure 5.15.: **Glade Scene**

Figure 5.16.: Scene Map of the **Glade Scene**

Experiment Part 2: Multi-Emotion Provoking Game Scenes

Two advanced game scenes were crafted for the second case study, and were tailored to the need of providing a natural, exciting and longer lasting game experience. This might for instance encourage endurance training. The objective is to provide a versatile method to design multi-emotion provoking game elements by postulating three composite strategies, which are further described in the appropriate scenes:

1. Repetition of the same type of provocation
2. An added obstacle in the pretext of the provocation
3. Providing several emotion provoking game elements

The **Frozensea Scene** includes distinct emotion provoking events. It was designed to evaluate the intensity in individual emotional reactions. The **Coin Collection Event** provides the participants with a challenging objective, similar to the **Parcours Scene** of the first experiment. The collection of four coins on an icy ground lowers a bridge to the finish line. During the player's ride through the scene environment a new designed **Jump Scare Event** is triggered. A couple of strange looking snowmen appear in front of the virtual bicycle with a loud and shocking sound. The scene design of multiple emotion provoking game elements provides the opportunity to directly compare the events and evaluate intensities in emotional responses.

Figure 5.17 displays the scene map of the **Frozensea Scene**.

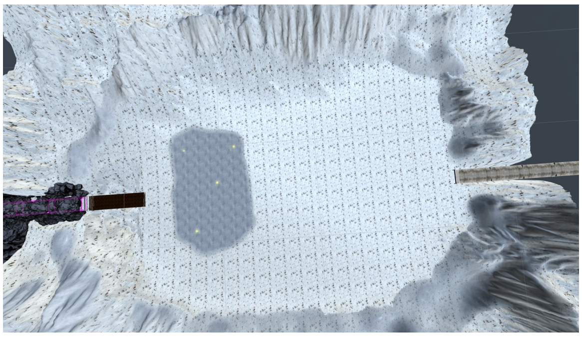


Figure 5.17.: Scene Map of the **Frozensea Scene**

The **Cliff Scene** includes stressful game elements, such as a narrow mountain path. Only careful steering allows the participants to reach the finish line. Usually many attempts are required to complete the task without falling off the cliff.

Figure A.8 displays the scene map of the **Cliff Scene**.

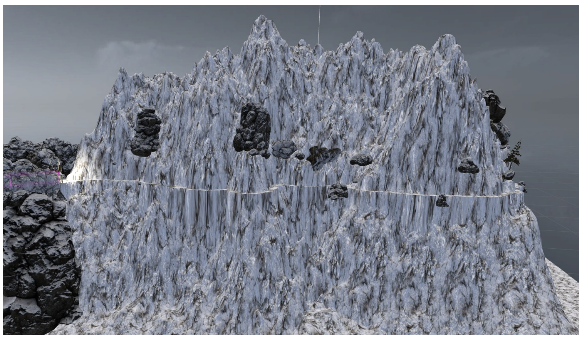


Figure 5.18.: Scene Map of the **Cliff Scene**

The **Falling Event** occurs in both the **Frozensea Scene** and **Cliff Scene**, thus the (1.) composite strategy was applied: repetition of the same type of provocation. In the following both scenes are described in more detail.

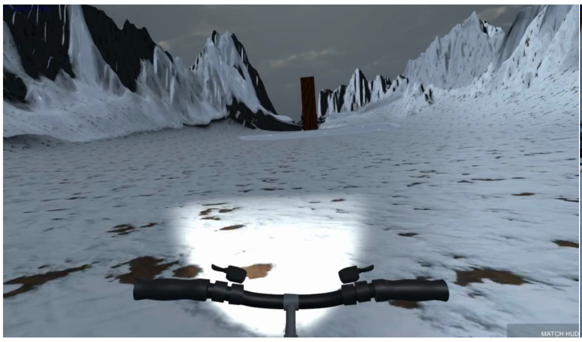
Frozensea Scene

The **Frozensea Scene** environment is full of ice and snow, as shown in Figure 5.19. Four coins have to be collected to lower a bridge to the finish line, as shown in Figure 5.19d. The coins are located on icy ground, and the virtual bike shows a very strange behaviour on the ground, as it simulates a real bike on ice with disabled brakes. The steering alignment is very challenging, and usually many attempts are needed to successfully collect all coins. The objective in this scene was similar to the **Parcours Scene** of the first experiment. The bridge, which is lowered after the successful collection of all coins, allows the participants to cross a canyon. If the player falls off the cliff edge or the bridge the **Falling Event** is triggered, accompanied by a scene restart. During the player's ride through the icy environment, a surprising **Snowmen Event** is triggered by the experimenter. Figure 5.19c shows the snowmen, which appear in front of the virtual bike, accompanied by the same horrible loud sound. The sound file is similar to the one from the **Jump Scare Event** in the **Forest Scene**. The duration of **Frozensea Scene** was approximately two minutes and 30 seconds, depending on the number of player died events.

Different events (**Falling Event**, **Snowmen Event**, **Coin Collection Event**) in the **Frozensea Scene** encourage to evaluation of the intensity of emotional reactions. Two composite strategies was applied in this scene. (3.) composite strategy of providing several emotion provoking game elements. The (2.) composite strategy of adding an obstacle in the pretext of a provocation was applied in this scene, as the player expects to have completed the task and does not anticipate another hidden challenge like the **Falling Event**.

Cliff Scene

In the **Cliff Scene** the virtual bicycle is placed high up an icy mountain, as presented in Figure 5.20. The only way to the finish line is a narrow pathway with a steep cliff. Loose rocks fall from the top of the mountain and cross the narrow pathway. These rocks sometimes hit



(a) Snowy Landscape

(b) Coin Collection Objective

(c) Surprising **Snowmen Event**

(d) Bridge to the Finish Line

Figure 5.19.: **Frozensea Scene** (Müller et al. (2017b), ©2017 IEEE)

the bike, which often corresponds with the **Falling Event**. The **Rock Hit Player Event** was perceived as very stressful as the player mostly falls off the cliff edge, accompanied by the **Falling Event**, which starts the scene again. The rocks are software controllable and can be triggered by the experimenter to maximise the participants' stressful experience. At the beginning of the scene a **Demo Rock Event** is triggered to showcase the possibility of being hit by a rock, although the demo rock does not hit the player. The duration of **Cliff Scene** was approximately eight minutes and 30 seconds, depending on the number of trials.

Very careful steering is required, and the player needs to ride very slowly to make it to the finish line despite the narrow pathway and sharp curves. At the end of the scene there is a narrow bridge to provoke fear, as shown in Figure 5.20. In case the player reaches the bridge the **Bridge Entered Event** is triggered. The whole scene provides the participants with very challenging and stressful objectives, therefore the experimenter offered a teleporting option after an excessive number of trials.

(a) Icy Mountain (b) Scary Bridge

Figure 5.20.: **Cliff Scene** (Müller et al. (2017b), ©2017 IEEE)

In the **Cliff Scene** the (3.) composite strategy of providing several emotion provoking game events was applied, as it includes the **Rock Hit Player Event**, the **Falling Event**, the **Demo Rock Event** and the **Bridge Entered Event**.

Experiment Part 3: Physical Stress and Emotional Provocation

The third part of the experiment was designed to combine physical effort with emotion provoking game elements. First the participants had to accomplish the **Treehouse Scene**, and the ascent of the virtual path was set to a high value. Subsequently the established **Mountain Scene** was selected to provoke surprise or fear through the **Spider Attack Event**.

The physical resistance of the pedals was set to a low maximum of 100 Watt for the half of the group. For the rest the maximum was set to 150 Watt, which means it was set related to the ascent of the scenes. These participants had to make a substantial effort to climb the hill with the virtual bike. Furthermore, the experimenter explained that the participants have to choose a high gear and start up the hill at a high speed to reach the finish line, as this is not obvious for a real bike.

The third part of the experiment was designed to evaluate assumed influences of physical effort on surprising events, and to provide more insights of physical stress combined with emotional provocations.

Treehouse Scene

The **Treehouse Scene** consists of a narrow wooden path, which the player has to follow to reach the finish line. The wooden path is designed as a configurable ascent and shows the enhancement of the new physic module (Requirement 4.1.3). In the presented experiment the ascent was set to a high value, as shown in Figure 5.21. It required much physical effort for half of the participants, as the physical pedal resistance increased in relation to the virtual slope. A maximum of 150 Watt was provided by the exercise machine and the participants had to use higher gears and ride up the hill very fast to make it to the finish line. The duration of **Treehouse Scene** was approximately one minute and 30 seconds. Figure 5.22 displays the scene map of the **Treehouse Scene**.

Figure 5.21.: **Treehouse Scene** (Müller et al. (2017b), ©2017 IEEE)

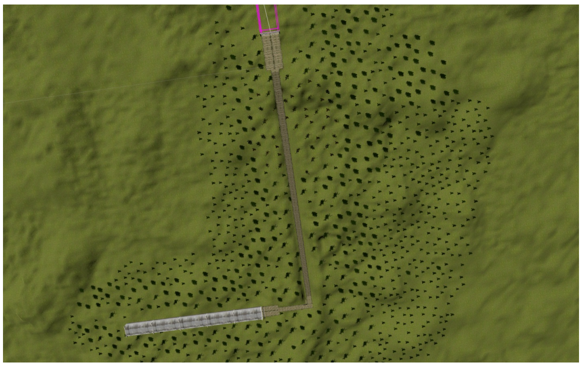


Figure 5.22.: Scene Map of the **Treehouse Scene**

Mountain Scene

The **Mountain Scene** was the second scene that required a physical effort from half of the group. Figure 4.7 illustrates that the design was similar to the **Mountain Scene** in the first case study. Although no unexpected walls merge from the ground in this experiment, as this event did not show reliable results. The number of spiders is configurable and was enhanced in the second case study to ensure the occurrence of a **Spider Attack Event**, as shown in Figure 5.23. The duration of **Mountain Scene** was similar to the duration in the first experiment. Figure 5.24 displays the scene map of the **Mountain Scene**.

Figure 5.23.: **Spider Attack Event** in the **Mountain Scene** (Müller et al. (2017b), ©2017 IEEE)

Figure 5.24.: Scene Map of the **Mountain Scene**

5.2. Data Analysis Method

In the second case study the system was enhanced by facial expression recognition, which provides the data in soft real-time (Requirement 2.5). The software Emotient from iMotions[1]

[1] https://imotions.com/emotient/

and Affectiva[2] was integrated for this experiment, due to the findings of Bernin et al. (2017). Affectiva provides probability values for anger, joy, surprise, sad, disgust, fear, contempt, smirk and smile. Emotient labels the emotions anger, joy, surprise, sadness, disgust, fear, contempt, confusion, frustration and neutral. The presented data analysis is limited to Emotient, which is the successor of CERT. It provides evidence values for frustration, as the experimental design includes the provocation of frustration.

The output provided by Emotient represents a distance to the hyperplane of the SVM and represents evidence value. Thus the data need to be normalised to probability values, as described in Bernin et al. (2017). The applied equation for the transformation is presented in 5.1.

$$p(x) = \frac{1}{1 + 10^{-1*x}} \tag{5.1}$$

The data analysis procedure was based on the findings of the first experiment, as described in Section 4.4.1: an event-based evaluation with an analysis window of 0.5 seconds before and 3.5 seconds after an event occurred.

The approach of an application specific clustering of emotional reactions in facial expressions was proposed by Bernin et al. (2017). Based on these findings, this work presents a scene specific clustering of emotions, as the recognition of more than one emotion is acceptable for many game elements. For the recognition of an emotional response to the **Jump Scare Event** for instance surprise is as acceptable as fear, as shown in Equation 5.2.

[2] https://www.affectiva.com/

The physiological data analysis in the second case study focuses on the evaluation of the EDA signal, as this sensor showed the most promising results in the first case study. The event-based analysis is further described in Section 4.4.2. Due to these findings, and based on an initial check of the new data, an analysis window from one second before and nine seconds after event occurrence was chosen. The applied EDA sensor in this experiment acquired a range of 0-25 μS.

In this work the results of EDA and facial expression recognition (FER) are combined to increase the successful recognition rate of individual emotional reactions. In detail both modalities are logically combined with an OR-operation (disjunction), as one sensor is sufficient for successful recognition, as shown in Equation 5.2.

In this experiment one sensor is sufficient for successful recognition, which would not be the case if the equation would include an AND-operation. If the participant for instance looks sideways and therefore no facial expression analysis is possible, but the EDA sensor provides clear data, it is a successful provocation, as some of the events come with a horrible sound that usually provokes a peak in EDA data. In addition, the EDA sensor might provide noisy data; in case there is a clear response in facial expression that is a successful recognition.

$$p(x) = A \vee B \tag{5.2}$$

5.3. Experimental Results

The presented case study analysed facial expressions, physiological data and actions of 25 participants. Two of the participants were very nervous during the experimental procedure, which might have influenced their physiological responses. The applied sensor could not measure their reactions in some of the scenes, as the value was above 25μS and thus outside of the range provided by the sensor. These EDA data were excluded from the analysis. For the other 23 participants stable results were obtained. In some of the scenes an error occurred due technical reasons, which caused a data loss. Thus in some scenes the emotion sensor data of the first proband are missing, resulting in n=24 for single events.

The analysis of facial expressions requires sufficient illumination of the participants' faces. In the second case study the acquired LUX values ranged from a minimum of 297 up to 370, with an average of 339 and a median of 335.

The participants finished a training scene before the three experimental tasks started. Thus they were familiar with the controls in the first evaluated scene. For the **Fork Scene** no results are presented, as this scene does not include any emotion provoking game elements.

5.3.1. Participants' Profile

In the second experiment the number of participants was raised to 25, ten males and fifteen females. Their ages ranged from 18 to 51 years with an average age of 29. It is important for the physiological data analysis, that the participants are not under the influence of any medication or drugs, which was confirmed by all of them. On average the participants were

1.73m tall, with a range from 1.56m to 1.96m. The saddle was adjusted to the participants' respective height to avoid distractions due to discomfort. Four participants had been patients of a medical rehabilitation clinic in the past, but none of them mentioned a knee or leg issue.

Fifteen participants said to exercise regularly, two exercised more than five hours per week and four did not exercise at all, or up to one hour per week. Two participants said to use a bicycle as a device for competitive sports, whereas two claimed to never use a bicycle. Sixteen stated they did bicycle tours from time-to-time, and ten to use the bike as means of transport. Five participants did physical exercises on the day of the experiment.

The preliminary studies showed that game experience might affect the perception of game elements, thus they were asked about their gaming behaviour; nine participants said to play video games at least two hours per week. The participants were asked to rate their gaming abilities on a scale from 0 to 5. Five of them stated zero and only one chose five. Eleven of the participants had no experience with new interaction technologies, such as Wii Controller, Kinect, Wii Balance Board, steering wheels, virtual reality or augmented reality.

Twenty-four participants gave their consent to the disclosure of their physiological data and questionnaire results to other researchers. Fifteen gave further consent to disclose their video data to other researchers.

5.3.2. Emotion Assessment

The detailed self-assessment is displayed in Appendix B. The observer-assessment is displayed in Appendix C.

5.3.3. Experiment Part 1: Emotional Journey Concept

The novel concept of an **Emotional Journey** was the first part of the experiment. After the participants finished the **Journey** the experimenter conducted a detailed interview, which included several questionnaires regarding the game concept and their experience. In addition, they were asked for a comprehensive self-assessment of all relevant game elements.

To evaluate the presented concept of an **Emotional Journey**, the participants were asked about their ability to recognise and reproduce the in-scene provocation. This qualitative analysis resulted in an overall recognition rate of 89.3% of the **Emotional Journey**. Each participant recognised the aim of the **Jump Scare Event** in the **Forest Scene**, to provoke fear or surprise, resulting in a recognition rate of 100%. The **Challenge Scene** was equally successful. All the participants claimed that it aimed at provoking frustration or ambition, which resulted in a recognition rate of 100%. The **Teddy Scene** and the **Glade Scene** were designed to provoke joy. This provocation was recognised by 68% of the participants, as the tailored game scenes had a more adverse effect than expected. In the detailed self-assessment, many participants stated they expected harm from the supposedly funny rabbit. Moreover, many participants stated that killing cute teddy bears made the feel uncomfortable. These results support the hypothesis about the necessity of fast and easy testing of game elements due to probably highly individual reactions (Müller et al. 2015).

The individual emotional responses in facial expressions and EDA data were analysed for all scenes of the **Emotional Journey**, and detailed results are presented in the following.

Forest Scene

The **Jump Scare Event** occurred three times per participant, except for two participants, where it was only triggered only twice (n=70). The results confirmed the hypothesis that the first occurrence will be perceived as the most frightening. Facial expression analysis showed a recognition rate of 62.5% for fear and 58.3% for surprise. The EDA data analysis displayed peaks in 91.7% of the cases, so the EDA is more reliable in this scene. For the second event the EDA rate decreased to 75%, fear to 45.8% and surprise to 12.5%. The flattened response of EDA data and fear probability is illustrated in Figure 5.25, and EDA data in combination with surprise in Figure 5.26. The small peaks at the beginning of the scene are related to the dark environment and not to a specific game event. Figure 5.25 also shows that for this event, the response in facial expressions appears much faster than in the EDA signal, which confirms the defined analysis windows for physiological data and facial expression recognition (FER) data.

In this scene the recognition of fear is as acceptable as surprise, thus a scene specific clustering of these two emotions was made, resulting in a combined FER value which is presented in Table 5.1. The recognition rate increased to 75% for the first event and to 50% for the second and the third event. For all events the recognition rate was enhanced to 58.6%, by the scene specific clustering.

Following the scene specific clustering the FER results were combined with the results of the EDA data analysis by a disjunction, as the successful recognition of one sensor is sufficient. This procedure enhanced the recognition rate to 95.8% for the first and the second event, as presented in 5.2. The overall recognition rate increased to 87.1%. The beneficial effect of a combination of these modalities in displayed in Figure 5.26. The participants displayed

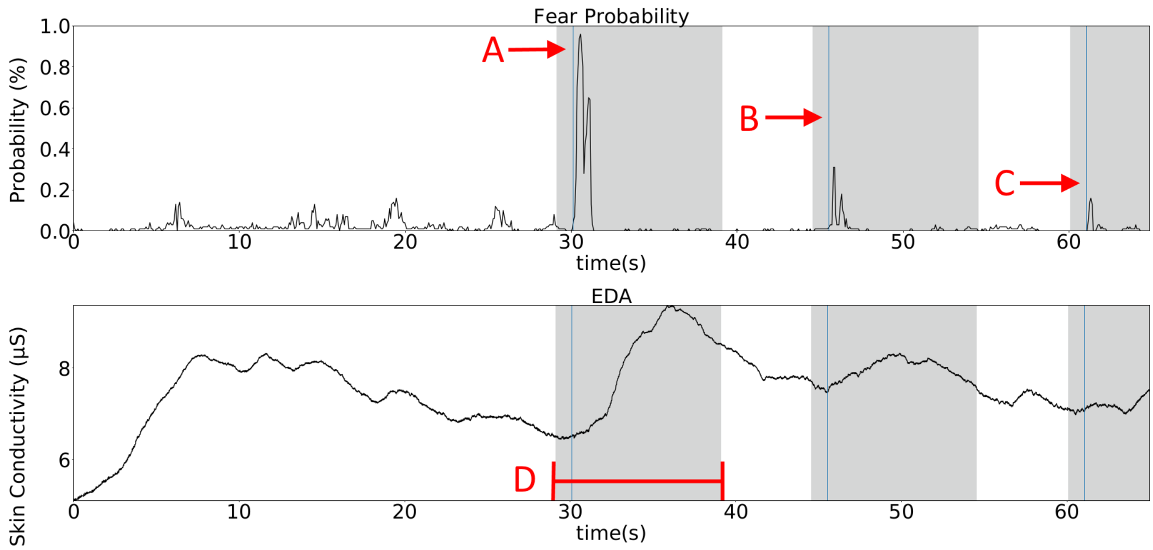


Figure 5.25.: Typical time series for the Jump Scare Event in the Forest Scene illustrating the flattening effect. The data consist of a. the facial expression probability of fear and b. the skin conductivity (EDA) raw data. A, B, C: Jump Scare Event 1, 2, 3 D: EDA Analysis Window (Müller et al. (2017a), ©2017 IEEE)

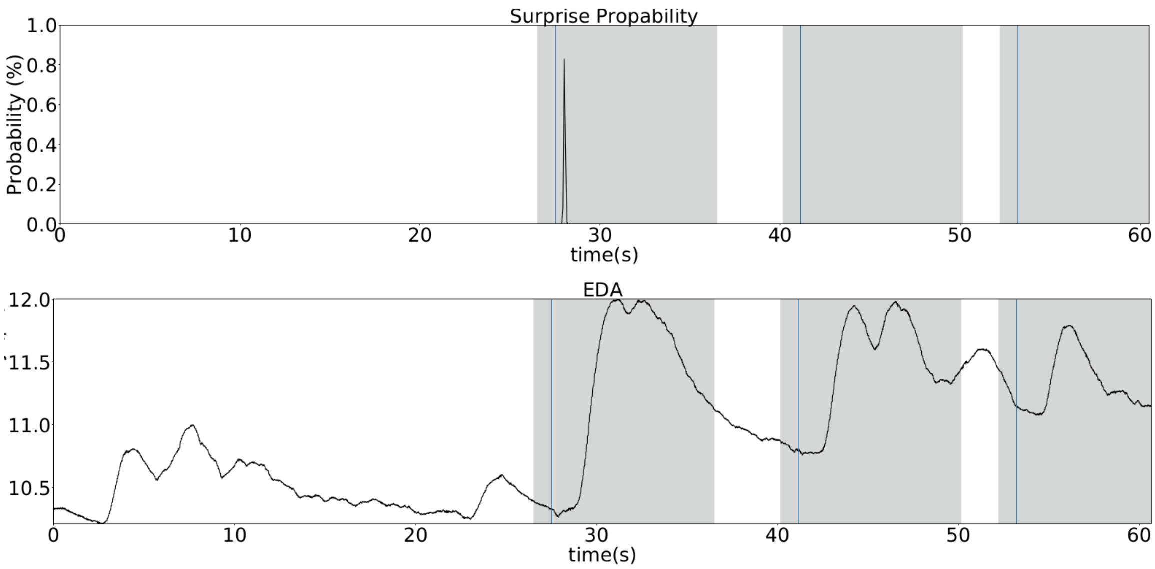


Figure 5.26.: Typical time series for the Jump Scare Event in the Forest Scene. The data consist of a. the facial expression probability of surprise and b. the skin conductivity (EDA) raw data.

Event Occurrence	FER Surprise	FER Fear	FER Combined
1	58.3	62.5	75
2	12.5	45.8	50
3	13.63	50	50
All	28.6	52.9	58.6

Table 5.1.: Occurrence of three subsequent events in combination of two FER Channel in the *Forest Scene*, Recognition Rates in % (Müller et al. (2017a), ©2017 IEEE)

Event Occurrence	EDA	Combined FER / EDA
1	91.7	95.8
2	75	95.8
3	45.5	68.2
All	71.4	87.1

Table 5.2.: Event occurrence-based combination of the two modalities FER and EDA in the *Forest Scene*, Recognition Rates in % (Müller et al. (2017a), ©2017 IEEE)

surprise only at the occurrence of the first event, but the EDA signal showed peaks at the second and the third event as well.

Explosion Scene

The analysis of EDA data showed peaks of 76.1% in the **Explosion Event** (n=71). This event was triggered multiple times per participant by the experimenter to maximise the surprising or fearful effect. The FER data analysis displayed a recognition rate of 46.5% for fear and 12.7% for surprise. The scene specific clustering of these two emotions resulted in a recognition rate of 47.9%. Table 5.3 provides an overview of the scene specific clustering for all scenes. The combination of EDA and FER analysis is illustrated in Table 5.4. In the **Explosion Scene** the recognition rate increased to 78.9%.

Scene	FER 1 Emotion	FER 2 Emotion	FER 1	FER 2	Combined FER 1 / FER 2
Forest	Fear	Surprise	52.9	28.6	58.6
Explosion	Fear	Surprise	46.5	12.7	47.9
Challenge	Frustration	Joy	24.7	90.6	95.3
Downhill	Frustration	Joy	30.6	84.7	86.7
Teddy	Joy		58.3		58.3
Glade	Joy		42.9		42.9
Mountain	Fear	Surprise	50.0	33.3	54.2

Table 5.3.: Scene specific Clustering of FER Channels, Recognition Rates in % (Müller et al. (2017a), ©2017 IEEE)

Scene	FER	EDA	Combined FER / EDA
Forest	58.6	71.4	87.1
Explosion	47.9	76.1	78.9
Challenge	95.3	84.7	96.5
Downhill	86.7	80.6	99
Teddy	58.3	91.7	100
Glade	42.9	95.2	100
Mountain	54.2	83.3	87.5

Table 5.4.: Two Modalities Combined: FER and EDA, Recognition Rates in % (Müller et al. (2017a), ©2017 IEEE)

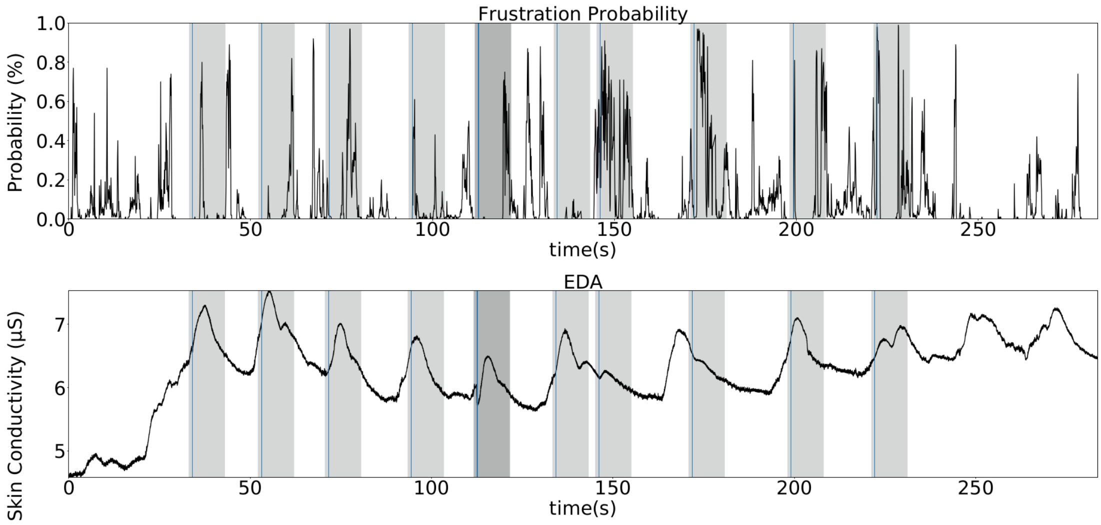


Figure 5.27.: Typical time series for the **Falling Event** in the **Challenge Scene**. The data consist of a. the facial expression probability of frustration and b. the skin conductivity (EDA) raw data. (Müller et al. (2017a), ©2017 IEEE)

Challenge Scene

The **Challenge Scene** includes a multiple number of **Falling Events**, depending on the participants' abilities to control the virtual bike and the software adjustment of repelling the player from landing on the other side automatically. Fortunately, some of the participants reached the other side despite this effect. The FER analysis results showed that frustration was detected in 24.7% of the events (n=85). For the reason described in Section 4.5.3, joy was also analysed and was displayed in 90.6% of the events. The recognition rate was enhanced to 95.3% by the scene specific clustering. The EDA data analysis showed peaks in 84.7% of the events and the combination of FER and EDA analysis increased the rate to 96.5%. Figure 5.27 and 5.28 show the combination of sample EDA raw data and FER frustration probability.

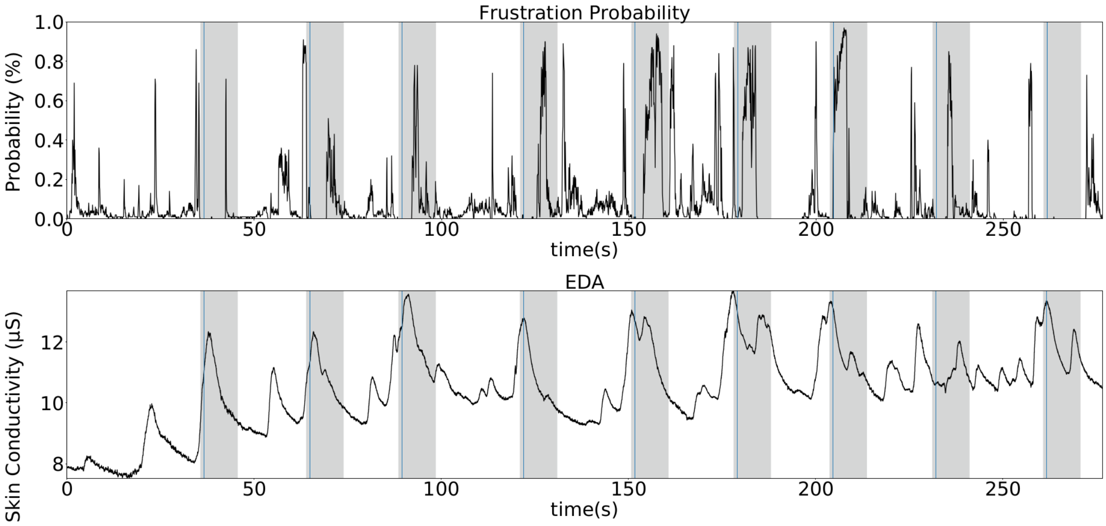


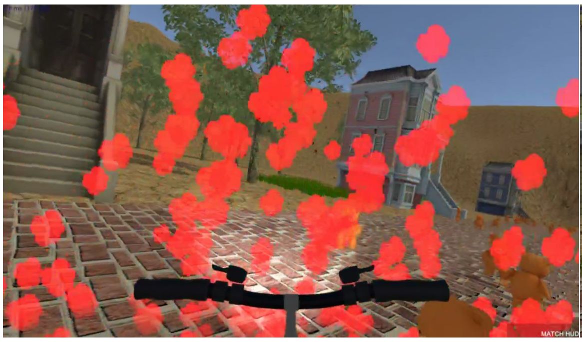
Figure 5.28.: Typical time series for the **Falling Event** in the **Challenge Scene**. The data consist of a. the facial expression probability of frustration and b. the skin conductivity (EDA) raw data.

Downhill Scene

In the **Downhill Scene** frustration was found in 30.6% of the **Falling Event**s (n=98) and joy in 84.7%. The experimenter influenced the players' controls by giving a speed boost in front of the narrow curve to maximise the frustrating effect, similar to the repel effect of the **Challenge Scene**, resulting in multiple attempts and **Falling Event**s. The scene specific clustering increased the number to 86.7%. The analysis results of the EDA data showed a recognition rate of 80.6%. The combination of FER and EDA enhanced the recognition to 99%. An interesting finding in the data analysis was that contempt had a very high detection rate of 82.6%.

Teddy Scene

In the **Teddy Scene** 58.3% of the participants displayed joy by seeing the cute teddy bears for the first time (n=24) at the **Scene Entered Event**. The EDA data analysis showed peaks in 91.7% of the events. The combination of EDA and FER enhanced the rate to 100%. An interesting finding in this scene was that fear was detected in 45.8% of cases. An explanation of that effect was given by the participants during the interview or questionnaire session after the **Emotional Journey** part of the experiment. The participants stated to feel uncomfortable with running over the teddy bears, but there were so many of them, that it was almost impossible to avoid all. Furthermore, the explosion effect following a **Teddy Hit Event**, as shown in Figure 5.29a, and the bodies of the teddy bears on the floor, as shown in Figure 5.29b, was not described as funny, as they were in preliminary studies.



(a) Explosion of a Teddy Bear

(b) Remaining Body of a Teddy Bear

Figure 5.29.: **Teddy Hit Event** in the **Teddy Scene**

Glade Scene

The **Glade Scene** was presented to many of the participants, as the provocation of joy in the **Teddy Scene** was less successful, resulting in an event occurrence of 21. The results

Favourite Scene	Number of Votes
Cliff Scene	7
Frozensea Scene	5
Forest Scene	4
Glade Scene	4
Teddy Scene	3
Challenge Scene	2
Mountain Scene	1

Table 5.5.: Favourite Scene Ranking of all Scenes

of seeing the funny rabbit in the **Glade Scene** was similar to the **Teddy Scene**. The EDA analysis showed peaks in 95.2%, while joy was detected in 42.9%. The combination of both increased the detection rate to 100%. An interesting finding was, that the participants stated that they felt scared by the funny big rabbit running around in the scene. They mentioned that they were expecting harm by the rabbit. A fear probability rate of 52.4% supports these statements.

5.3.4. Experiment Part 2: Multi-Emotion Provoking Game Scenes

Two multi-emotion provoking game scenes were designed for the second case study. The user perception of those scenes was much higher compared to single-provoking scenes, as both were mentioned as their favourite scene by most of the participants. Table 5.5 illustrates the favourite scene ranking for the whole experiment. Seven participants voted for the **Cliff Scene** and five for the **Frozensea Scene**.

The multi-emotion provoking game scenes were designed to for instance evaluate different intensities in emotional reactions to particular events in one game scene. Figure 5.30 displays the combination of the **Rock Hit Player Event** and the **Bridge Entered Event Event**, and thus illustrates that both events can be directly compared in the analysis. The analysis of composite strategy (3. Providing several emotion provoking game objectives) in the **Cliff Scene** shows that the **Rock Hit Player Event** provokes the most stressful responses. The **Falling Event** and the narrow bridge did not trigger the same reliable results. The analysis of composite strategy (3. Providing several emotion provoking game objectives) **Frozensea Scene** illustrates that the **Falling Event** provokes more intense reactions than the **Snowmen Event** or the **Coin Collection Event**.

The analysis of composite strategy (1. Repetition of the same type of provocation), displays that the **Falling Event** provoked stronger reactions in the **Frozensea Scene**, due to composite strategy (2. An added obstacle in the pretext of a provocation). Different intensities of emotional reactions to similar game events in different game scenes imply that it is important to consider the whole game context and all emotions in the order of provocation. Hereafter the detailed scene results of the multi-emotion provoking scenes (**Frozensea Scene** and the **Cliff Scene**) are presented.

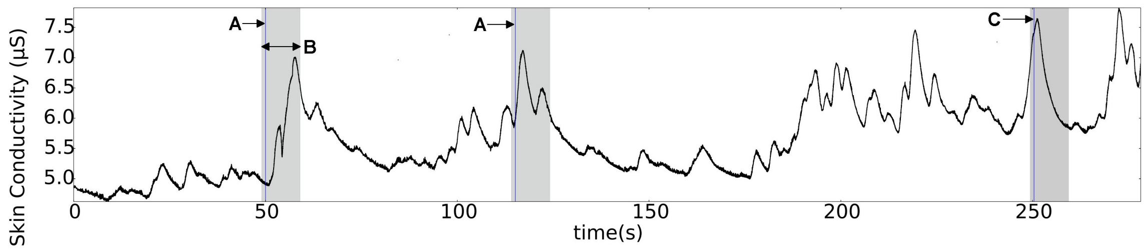


Figure 5.30.: Typical time series of the **Cliff Scene** consists of the EDA data, A: **Rock Hit Player Event**, B: Analysis Window, C: **Bridge Entered Event Event** (Müller et al. (2017b), ©2017 IEEE)

Frozensea Scene

The **Snowmen Event** in the **Frozensea Scene** provoked fear in 39.5% and surprise in 17.1% of the events (n=76). As both emotions are acceptable in this case, the scene specific clustering led to 44.7%, as shown in Table 5.6. The analysis of EDA data showed peaks in 84.2% of the cases. The combination of EDA and FER analysis increased the rate to 90.8%, as shown in Table 5.2. Figure 5.31 shows a typical time series of EDA raw data of the **Snowmen Event** in the **Frozensea Scene**.

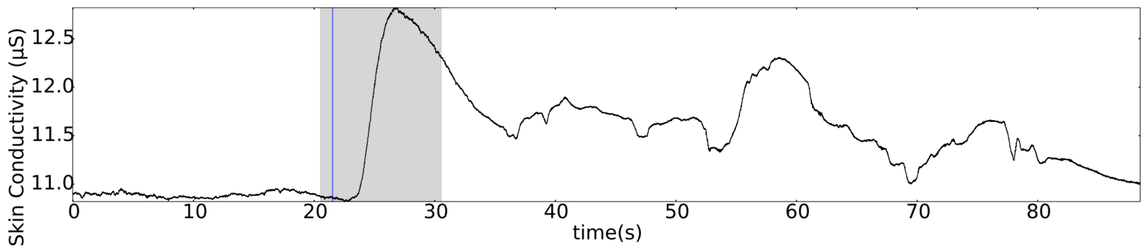


Figure 5.31.: Typical time series **Frozensea Scene** for the **Snowmen Event** (Müller et al. (2017b), ©2017 IEEE)

Joy was displayed in 43.3% of the **Coin Collection Events** (n=164). The scene started again in case the participants fell off the cliff, resulting in a high number of events. Eleven of the participants collected all coins in the first trial. Ten needed two attempts to make it, one needed three and one person needed four attempts to reach the finish line. EDA data showed peaks in 62.2% of the **Coin Collection Events**. The number could be increased to 72% by combining EDA and FER analysis.

Frustration was displayed in 25% and joy in 87.5% of the **Falling Events** (n=16). The scene specific clustering resulted in 93.8%. The combination of EDA and FER analysis increased the detection rate to 100%. The **Falling Event** was perceived as very frustrating, although the repetition of this event is much less than in other scenes. It takes a long time to finish the

Scene	Event	FER 1 Emotion	FER 2 Emotion	FER 1	FER 2	Combined FER 1 / FER 2
Frozensea	Snowmen	Fear	Surprise	39.5	17.1	44.7
Frozensea	Coin Collection	Joy		43.3		43.3
Frozensea	Snowmen	Frustration	Joy	25.0	87.5	93.8
Cliff	Falling	Frustration	Joy	25.0	65.9	70.5
Cliff	Bridge Entered	Fear	Surprise	36.4	13.6	45.5
Cliff	Rock Hit Player	Fear	Surprise	56.3	18.8	56.3
Cliff	Demo Rock	Fear	Surprise	37.5	12.5	41.7

Table 5.6.: Scene specific Clustering of FER Channels for the **Frozensea Scene** and the **Cliff Scene**, Recognition Rates in % (Müller et al. (2017b), ©2017 IEEE)

Scene	Event Name	FER	EDA	Combined FER / EDA
Frozensea	Snowmen	44.7	84.2	90.8
Frozensea	Coin Collection	43.3	62.2	72.0
Frozensea	Falling	93.8	87.5	100
Cliff	Falling	70.5	79.5	90.9
Cliff	Bridge Entered	45.5	77.3	90.9
Cliff	Rock Hit Player	56.3	87.5	93.8
Cliff	Demo Rock	41.7	62.5	75

Table 5.7.: Two Modalities Combined: FER and EDA for the **Frozensea Scene** and the **Cliff Scene**, Recognition Rates in % (Müller et al. (2017b), ©2017 IEEE)

collection of all the coins. This might lead to a much stronger perception of frustration, even with minor occurrences than in other scenes.

Cliff Scene

In the **Cliff Scene** frustration was provoked in 25.0% by the **Falling Event** (n=44). Joy was displayed in 65.9%. The rate increased by the scene specific clustering to 70.5%, as the EDA showed peaks in 79.5% of the events. FER and EDA combined resulted in a detection rate of 90.9%.

Fear was displayed in the facial expressions in 36.4% and surprise in 13.6% by the **Bridge Entered Event** (n=22). The combination resulted in 45.5%, while peaks were displayed in 77.3% of the EDA data. FER and EDA combined resulted in 90.9%.

The **Rock Hit Player Event** provoked fear in 56.3% and surprise in 18.8% (n=16). The applied scene specific clustering led to 56.3%. The EDA data analysis resulted in 87.5%. FER and EDA combined resulted in 93.8%.

Fear was displayed in the facial expressions in 37.5% and surprise in 12.5% of the **Demo Rock Triggered Events** (n=24). The combination increased it to 41.7%. Peaks were found in 62.5% of the events. FER and EDA combined increased the detection rate to 75%.

Two of the participants asked for the option to be teleported to the finish line. Many attempts were necessary, but the rest wanted to make it on their own. This indicates a compelling and exciting effect of the entertainment.

5.3.5. Experiment Part 3: Physical Stress and Emotional Provocation

The results of the case study indicate, that the physical effort of the **Treehouse Scene** and the **Mountain Scene** did not have a huge impact on the emotional reactions. This conclusion can be made on a qualitative level. An experimental design with more participants would be required for a quantified assertion. In particular the combination of FER with EDA data analysis enhanced the detection rate to 87.5% for the **Spider Attack Event**, confirming that there was no remarkable influence of physical effort on the emotion provocation.

An interesting finding is that even participants without a high physical pedal resistance stated that they perceived a challenge or strain by riding up the visual ascent in both scenes. Below a detailed scene evaluation is provided.

Treehouse Scene

The **Treehouse Scene** did not contain any emotion provoking game elements, therefore none were evaluated. The scene was tailored to provide a physical effort, by presenting an extreme ascent in the virtual environment.

Mountain Scene

In the **Mountain Scene** the **Spider Attack Event** provoked fear in 50% and surprise in 33.3% of the facial expressions. Figure 5.32 illustrates a typical time series for the **Spider Attack Event**, consisting of a. the facial expression probability of fear and b. the EDA raw data. A typical time series is presented in Figure 5.33. The scene specific clustering of these two emotions increased the detection rate to 54.2%. The EDA raised peaks in 83.3% of the events (n=24). Figure 5.33 and Figure 5.32 illustrate the beneficial effect of combining the modalities FER and EDA data, which furthermore enhances the detection rate to 87.5%.

5.4. Conclusion

In the presented nonlinear gameplay concept, individual emotional reactions can lead to different and individual experiences, as the emotion provoking game design was enhanced

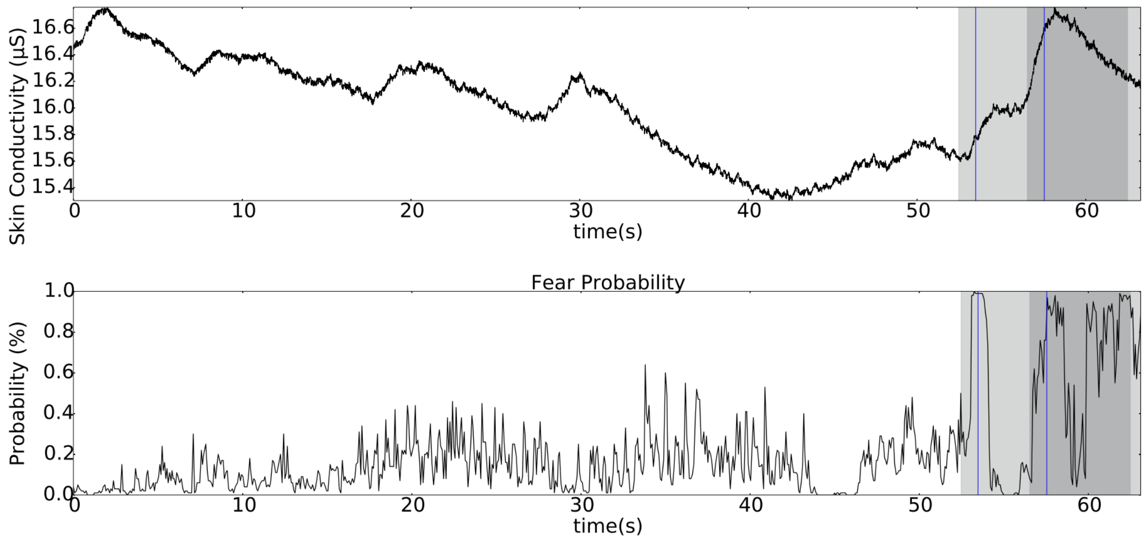


Figure 5.32.: Typical time series for the **Spider Attack Event** in the **Mountain Scene**. The data consist of a. the facial expression probability of fear and b. the skin conductivity (EDA) raw data. (Müller et al. (2017b), ©2017 IEEE)

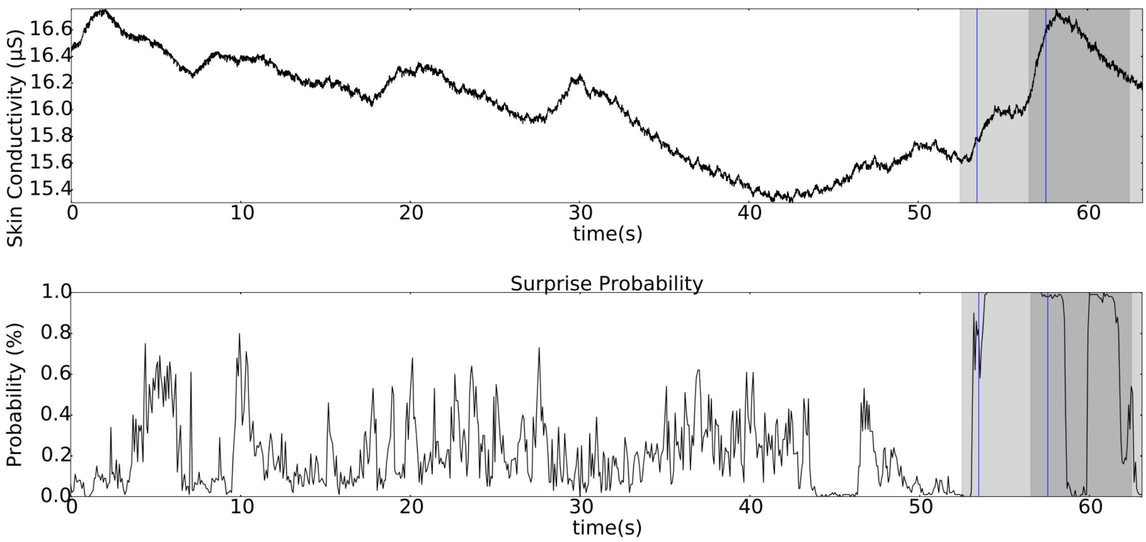


Figure 5.33.: Typical time series for the **Spider Attack Event** in the **Mountain Scene**. The data consist of a. the facial expression probability of surprise and b. the skin conductivity (EDA) raw data.

with reactive system response capabilities. Furthermore, the physical exergame controller was enhanced by a gearshift and a brake to facilitate more thrilling events and enable a more natural and immersive cycling game experience.

In this chapter a case study with 25 participants was described to showcase the ability of the system to dynamically change its game play in response to users' emotional reactions. The introduced novel **Emotional Journey** concept for affective entertainment provides a dynamic and adaptive story path for players, based on their emotional reactions. The empirical results illustrated the success by a recognition rate of 100% for two of the three parts. The participants were not able to reproduce the in-scene provocation of the third part reliably. Many of them said to be scared by the idea of driving over the cute teddy bears, and the designed explosion of red cotton wool only provoked joy a few times. The lesser success of the third part might be related to the scene order. Most of the participants stated that they were expecting danger from the funny rabbit after what happened to the teddy bears. A further interesting finding for designing emotion provoking game elements was the flattened emotional responses by the repetition of the **Jump Scare Event** in the **Forest Scene**. The detection rate for the first and the second event was 95.8%, while the third only reached 68.2%. Furthermore, the provocation and recognition of emotions such as frustration make the presented findings highly interesting for building emotional adaptive learning applications.

The analysis method of facial expression that was designed for the first case study was enhanced by soft real-time abilities in the second case study. Moreover, the introduced scene specific clustering of emotions increased the emotion recognition rates. The EDA analysis was enhanced by wireless devices to avoid the detriments of the cable-based approach. The combination of the FER clustering with the EDA analysis that was presented in this chapter resulted in more robust emotion recognition rates. The presented method could increase the detection rate of the **Falling Event** in the **Frozensea Scene** for instance to 100%.

This chapter further introduced new entertaining content for fitness applications by multi-emotion provoking game scenes. The FER and EDA data analyses resulted in interesting findings, as they showcased individual differences in emotional responses and intensities in reactions to distinct game elements. The three introduced composite strategies to effectively integrate stress and emotion provocation in game scenes proved to be useful methods to tailor new entertaining content for fitness applications. The **Rock Hit Player Event** was perceived as the most stressful in the **Cliff Scene**, while the **Falling Event** in this scene did not provoke as reliable reactions as in other scenes. The presented findings emphasise the importance of evaluating multi-emotion provoking elements in affective games.

In this work physical stress was combined with emotional provocation, to provide an exciting but not too taxing or stressful training. The **Spider Attack Event** in the **Forest Scene** provoked emotional responses in most of the participants, even though a previous task included a high physical effort.

The presented exergame system enables motivational aspects, due to its possibility to enhance exercise endurance and intensity. Furthermore, the emotion provoking game scenes are chained together, resulting in a dynamic game experience. Overall, the results of the case study showed a positive response to the affective game experience by the participants.

In summary, it has been shown that it is possible to mix emotional provocation and physical stress, but it requires some design care in timing and story-telling elements. The novel concept of an **Emotional Journey** provides a dynamic and adaptive story path for a player based on emotional responses. Moreover, the presented analysis method, which combines the scene specific clustering of facial expressions with EDA data, results in more robust emotion recognition rates. The presented findings make this work a step toward more engaging and emotion responsive fitness applications.

6. Conclusion and Discussion

This chapter presents a thesis summary that contains experimental results and relevant findings for the research field of affective computing. The main contributions are illustrated and future perspectives are discussed.

6.1. Summary and Conclusion

In this work affective computing was applied in an exergame application, and new interaction methods were evaluated with this non-traditional human-computer interface. Affective computing methods are well studied in standard desktop scenarios, but not in exergaming applications. Two novel event-based emotion analysis methods were presented that can be transferred to other applications. Moreover, the concept of an Emotional Journey illustrated that it is possible to react to user emotions and provide a dynamic game play.

To summarise, this work enhanced a cycling exercise machine to act as a game controller and provide a controlled exergame environment. An ergometer was physically improved by steering capabilities, a gear shift and a brake to control a virtual bicycle. In a virtual fun-racer game the player had to ride a virtual bike through a crafted environment. The ergometer RPM was made software readable and can be set related to the virtual surrounding, e.g. riding up a hill. Furthermore, the game was designed to steer the participants in controlled emotional states. Thereby an emotion provoking exergame was introduced that is able to provoke single- and multi-emotions and also to support physical provocation.

This work also presented novel approaches for emotion analysis. A multimodal sensor data acquisition was designed, which applies a loosely coupled architecture to support an easy integration of new emotion sensors. A novel event-based analysis method for provoked facial expressions in cycling exergames was presented and enhanced by a novel event-based analysis method for EDA data. The analysis was enhanced by scene specific clustering and was made available in real-time to enable a dynamic system response. The affective game concept of a nonlinear game play was introduced: the Emotional Journey. It provides a system adaptation to a user's emotional reactions. The designed exergame system is a first step toward an exciting and motivating exercise that does not tax the user.

Two experiments were designed to evaluate the presented affective system. The first experiment focused on the emotional provocation of participants and its evaluation. For the second experiment, the setup was enhanced to enable a dynamic system response. Both conducted experiments provide a proof of concept for the defined research questions described in Section 2.6. Moreover, the second experiment generated a database of genuine emotional reactions, containing frontal video data, thermal video data, physiological data and personality questionnaires. Table 6.1 summarises the two experiments in comparison.

	Experiment A	Experiment B
Main Application Area	Game	Healthcare, fitness
Related fields	Learning or cockpit applications	Adaptive learning systems, biofeedback companion systems
Setup	Classical laboratory setup	Smart home environment
Controls	Steering, moving the virtual bike by pedalling, linking of physical pedal resistance and virtual ascent	+ Brake, gearshift, physical module
Participants	11	25
Database	No	Yes
Emotion assessment	Free choice, brief between scenes	Predefined list of emotions, brief between scenes and extensive at the end
Vision-based sensors	Kinect	+ Thermal camera
FER	CERT	Emotient, Affectiva
Physiological sensors	EDA, respiration, pulse, temperature change	+ ECG
Physiological system	BIOPAC, cable-based	Biosignal Plux, wireless
Aim of the experiment	Emotion provocation, emotion analysis, single-emotion game events	Emotion provocation of new scenes, physical effort, multi-emotion game scenes, dynamic game play through individual emotional reactions
Emotion analysis	FER analysis, physiological data analysis: EDA baseline, event-based peaks, respiration, temperature change	FER with scene specific clustering, EDA event-based peak detection, combined FER and EDA
Modus Operandi	Post-game analysis of provocation success, testing of game elements	Soft real-time enhancement for dynamic system control to provide a dynamic path of emotions (journey) based on individual emotional reactions, provide exciting but not too taxing experience

Table 6.1.: Summary of the Two Experiments

6.2. Main Contributions

The impact of this thesis was demonstrated in the peer-reviewed publications Müller et al. (2015), Müller et al. (2016), Müller et al. (2017a), and Müller et al. (2017b). The contributions are described in the following.

This work gained insights into related research of the highly interdisciplinary field of affective computing. It discussed various emotion theories that are applicable to affective systems and illustrated that discrete modelling of emotions is as acceptable as dimensional modelling. Moreover, it showed that exergames are a promising way of emotion provocation, and presented an overview of related affective system approaches and emotion recognition techniques. This work discussed that experiments are required to evaluate affective systems, thus it also described methodologies.

Chapter 3 presented the design and setup of a controlled exergame environment. An emotion provoking virtual cycling game was crafted and a physical exercise machine was enhanced to act as a game controller. The setup was designed with a loosely coupled architecture and provided a testbed for different algorithms and applications for emotion recognition and emotion provocation. An emotion provoking game was designed to steer participants in controlled emotional states. Single- and multi-emotion provoking game events were crafted, and a concept of a dynamic game play by means of emotional adaptation was introduced: the Emotional Journey.

An analysis method was developed to recognise the emotions of exercising participants through their facial expressions (Müller et al. 2015). It was improved through the analysis of physiological data (Müller et al. 2015). It was also enhanced by scene specific clustering and soft real-time abilities to enable a dynamic system response (Müller et al. 2017a; Müller et al. 2017b).

Two experiments were designed and conducted to evaluate the affective system. In the first case study, described in Chapter 4, eleven participants were invited to test single-emotion provoking game scenes. The results of the first experiment showed that the tailored game elements are able to provoke specific emotions. The results of the facial expression analysis illustrated that the **Teddy Hit Event** and the **Falling Event** provoked clear emotions. The **Teddy Hit Event** provoked joy in 80% and the **Falling Event** did the same in 90.9% of the participants. A more varied emotional response in type and probability was found for the **Coin Collection Event** and **Jump Scare Event**. The experiments demonstrated that the combination of physiological data analysis with facial expression analysis magnifies the potential for affective exergaming applications. The analysis of the EDA data resulted in a detection rate of 97% for the **Falling Event** and 91% for the **Jump Scare Event**.

The second experiment was designed to showcase that the system is able to dynamically adapt its game play based on the emotional reactions of the participants. New entertaining content was presented, which included multi-emotion provoking game scenes. It also mixed the emotional provocation with physical stress. An experiment was conducted with 25 participants, as described in Chapter 5. Thereby a database of genuine emotions was created. The introduced **Emotional Journey** concept provided a dynamic and adaptive story path based on emotional reactions. It showed promising results for the predefined emotional path of three emotions. Two of the three emotional provocation aims were reliably described in the retrospection, resulting in a rate of 100%.

The soft real-time analysis method of facial expressions and EDA data was applied, resulting in solid recognition rates. Scene specific clustering of emotions was used for the facial expression analysis. The new analysis method significantly increased the successful recognition rates. The detection rate of the **Falling Event** in the **Frozen Sea Scene** could be enhanced to 100% with the presented soft real-time multimodal emotion analysis method. The detection rate of the **Snowmen Event** in the **Frozensea Scene** was increased to 90.8%. The detection rate of the **Rock Hit Player Event** in the **Cliff Scene** to 93.8% and the rate of the **Falling Event** as well as the rate of **Bridge Entered Event** to 93.8%. The detection rate for the **Falling Event** in the **Challenge Scene** was increased to 96.5%. Moreover, the results showed a flattening effect of the **Jump Scare Event** in the **Forest Scene**. The detection rate for the first and the second time the event occurred was 95.8%, which decreased to 68.2% for the third occurrence.

In this work affective computing methods were combined with a HCI application. It showed that many affective computing techniques are already applicable in a laboratory setting, and can be utilised for a non-traditional setup: an exergame. Emotions were provoked by crafted game events, thereby a database was created containing non-acted emotional reactions in an exercising context.

This thesis closes the research gaps defined in 2.6, as the experiments showed that the system is able to provoke emotions by tailored game elements and to detect them by an event-based analysis method. Furthermore, the new paradigm of an Emotional Journey was presented. With this concept it has been shown that it is possible to add user emotions to create an adaptive story path.

6.2.1. Discussion of Research Questions and Objectives

The first research question that is answered in this work is (1) Can a testbed be provided for various affective applications to provoke specific emotions in an exercise context?

In this work a realistic testbed for various real-world affective applications was presented. It is designed as a controlled exergame environment to minimise irrelevant factors and to create reproducible results. It includes a multimodal sensor data acquisition system to provide a more complete view for emotion recognition. The system architecture is extensible and flexible. Different emotion sensors are included to enhance the emotion recognition rates.

Other related research often focuses solely on improving emotion detection algorithms. In this work emotion recognition is applied in an exergaming scenario, as it requires physical effort in order to progress the game. Furthermore, the emotional reactions of the participants are analysed, and the system is able to adapt the gameplay according to the expressed emotions.

The emotional provocation introduced in this work consists of single- and multi-emotion provoking gaming strategies for an enlarged study of diverse emotional provocations. Related research often applies casual games in their studies. In this work game scenes were designed and crafted to provoke specific emotional reactions.

The second research question addressed in this work was (2) Can existing emotion recognition methods be applied to support a reliable emotion recognition in an exergaming context?

In this work two novel event-based emotion analysis methods were introduced for higher emotion recognition precision in an exergaming context. Related research in the area is often based on desktop scenarios. The domain of an exergame differs significantly from a desktop scenario. The emotion recognition algorithms are influenced by the movement of the participant. Many facial expression recognition algorithms for instance have less successful recognition rates when the face moves.

The third research question addressed in this work was (3) How can a concept of a dynamic system response to emotional user reactions in the testbed be applied and evaluated?

This work introduced a new affective game concept of a nonlinear game play for an enhanced gaming experience. The Emotional Journey was designed to provide a more individualised experience by a runtime adaptation to emotional reactions.

The presented system was evaluated by means of two experiments. These experiments were designed and conducted to validate, test, and evaluate the proposed methods and concepts. The experiments were conducted with in summary 36 participants, and demonstrated that the proposed methods and concepts are successful. In addition they resulted in four peer-reviewed publications (Müller et al. 2015; Müller et al. 2016; Müller et al. 2017a; Müller et al. 2017b)

6.2.2. Limitations

The presented system allows to analyse various interesting aspects of affective computing technologies. The transferability for other applications has to be evaluated for every individual case. However, enhancements can be applied through iterative procedures, as it was shown in the second iteration of the PhD thesis. The flexible and extensible architecture allows to easily exchange parts of the system. A casual game controller for instance can be used to control the game. The game design allows to easily exchange the virtual bike asset. The physic module allows to for instance model a car in the same virtual environment. In this case the steering could be applied to a steering wheel. However, one limitation of such an application would be the lack of required effort to make progress in the game, which is one of the main advantages of the system.

The flexible and extensible architecture additionally allows to replace the ergometer with another exercise machine, for instance a treadmill. However, it would be difficult to provide changes in the physical effort of running up a virtual hill. Minimal requirements for other exercise machines are steering and speed, although this might except a slope related provocation.

6.3. Discussion and Future Perspectives

In this work a novel setup for affective computing experiments was presented. The system architecture enables an easy exchange of the application and the components. Additional emotion sensors can be integrated into the multimodal recognition system to improve recognition rates. The emotion provocation can be replaced by other stimuli or stressors. A casual game controller or another physical exercise machine can replace the ergometer. This flexible system architecture enables the transferability to different kinds of future experiments in affective computing or related research areas.

The presented multimodal analysis methods for provoked emotions can be generalised for various types of applications. Games for instance can be tested during their development, or stress can be evaluated in cockpit scenarios. A main advantage of these methods is their event-based design. Recognising the emotional response to a specific event enables the system to dynamically adapt its parameters. These novel methods empower the development of event-based communication concepts for future companion systems.

The system is a first step toward a fitness companion. The physiological data can be applied to control the exercise machine and the game environment. This would enable the retention of the physical effort for each user in an appropriate range, to provide a challenging but not too taxing training experience. Game events and emotional responses can be applied to learn user preferences in order to keep the player motivated; for some players it may be more interesting to complete challenging tasks, while others enjoy riding through peaceful environments. Machine learning approaches may be used to provide a trained and individualised user model, which would enable more individualised story-telling. This approach is a main advantage compared to the typical fitness tracker, which only provides biofeedback. An EmotionBike fitness companion would be able to adapt the exercise environment, instead of merely informing the users about their current fitness level. A fitness companion that recognises the users' needs, emotions and physiological states can adapt the exergame environment and provide a new level of individualised training experiences. This would yield a contribution to the health-related quality of life.

Bibliography

Abadi, Mojtaba Khomami et al. (2013). “Multimodal engagement classification for affective cinema”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 411–416.

Afzal, Shazia and Peter Robinson (2011). “Natural Affect Data: Collection and Annotation”. In: *New Perspectives on Affect and Learning Technologies*. Ed. by Rafael A. Calvo and Sidney K. D’Mello. New York, NY: Springer New York, pp. 55–70. ISBN: 978-1-4419-9625-1. DOI: 10.1007/978-1-4419-9625-1_5. URL: https://doi.org/10.1007/978-1-4419-9625-1_5.

Akhanda, Md Abu Baker Siddique, Shaon Md Foorkanul Islam, and Md Mostafizur Rahman (2014). “Monitoring the performance of computer user by analyzing physiological signals”. In: *Computer and Information Technology (ICCIT), 2013 16th International Conference on*. IEEE, pp. 120–125.

Al Khatib, Iyad et al. (2006). “A multiprocessor system-on-chip for real-time biomedical monitoring and analysis: architectural design space exploration”. In: *Proceedings of the 43rd annual Design Automation Conference*. ACM, pp. 125–130.

Alaoui-Ismaïli, Ouafe et al. (1997). “Basic emotions evoked by odorants: comparison between autonomic responses and self-evaluation”. In: *Physiology & Behavior* 62.4, pp. 713–720.

Alghowinem, Sharifa et al. (2013). “Head pose and movement analysis as an indicator of depression”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 283–288.

Almaev, Timur R and Michel F Valstar (2013). “Local gabor binary patterns from three orthogonal planes for automatic facial expression recognition”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 356–361.

Alyuz, Nese et al. (2016). “Semi-supervised model personalization for improved detection of learner's emotional engagement”. In: *Proceedings of the 18th ACM International Conference on Multimodal Interaction*. ACM, pp. 100–107.

Amini, Reza and Christine Lisetti (2013). “HapFACS: An open source API/software to generate FACS-based expressions for ECAs animation and for corpus generation”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 270–275.

Arnold, Magda B. (1960). *Emotion and Personality Vol 1 & 2*. Columbia University Press.

Asthana, Akshay et al. (2009). “Evaluating AAM fitting methods for facial expression recognition”. In: *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. IEEE, pp. 1–8.

Averill, J. R. (1980). “A Constructivist View of Emotion”. In: Emotion: Theory, Research and Experience. Academic Press, pp. 305–339.

Aviezer, Hillel et al. (2008). “Angry, disgusted, or afraid? Studies on the malleability of emotion perception”. In: *Psychological Science* 19.7, pp. 724–732.

Baggett, H Lane, Patrice G Saab, and Charles S Carver (1996). “Appraisal, coping, task performance, and cardiovascular responses during the evaluated speaking task”. In: *Personality and Social Psychology Bulletin* 22.5, pp. 483–494.

Barrett, Lisa Feldman et al. (2007). “The experience of emotion”. In: *Annual review of psychology* 58, p. 373.

Barros, Linda, Pilar Rodriguez, and Alvaro Ortigosa (2013). “Automatic classification of literature pieces by emotion detection: A study on quevedo’s poetry”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 141–146.

Bartlett, Marian Stewart et al. (1999). “Measuring facial expressions by computer image analysis”. In: *Psychophysiology* 36.2, pp. 253–263.

Bartlett, Marian Stewart et al. (2003). “A prototype for automatic recognition of spontaneous facial actions”. In: *Advances in neural information processing systems*. MIT; 1998, pp. 1295–1302.

Bartlett, Marian Stewart et al. (2006). “Automatic recognition of facial actions in spontaneous expressions.” In: *Journal of multimedia* 1.6, pp. 22–35.

Baumann, Timo and David Schlangen (2012). “The INPROTK 2012 release”. In: *NAACL-HLT Workshop on Future Directions and Needs in the Spoken Dialog Community: Tools and Data*. Association for Computational Linguistics, pp. 29–32.

Becker, Christian (2003). “Simulation der Emotionsdynamik eines künstlichen humanoiden Agenten”. MA thesis. University of Bielefeld. URL: http://www.becker-asano.de/DA_Komplett.pdf.

Becker, Christian and Ipke Wachsmuth (2006). “Modeling primary and secondary emotions for a believable communication agent”. In: *Proceedings of the 1st Workshop on Emotion and Computing*, pp. 31–34.

Becker, Christian et al. (2005). “Evaluating affective feedback of the 3D agent max in a competitive cards game”. In: *International Conference on Affective Computing and Intelligent Interaction*. Springer, pp. 466–473.

Becker-Asano, Christian (2008). *WASABI: Affect simulation for agents with believable interactivity*. Vol. 319. IOS Press.

– (2011). “Affective computing combined with android science”. In: *KI-Künstliche Intelligenz* 25.3, pp. 245–250.

Becker-Asano, Christian and Hiroshi Ishiguro (2011). “Evaluating facial displays of emotion for the android robot Geminoid F”. In: *Affective Computational Intelligence (WACI), 2011 IEEE Workshop on*. IEEE, pp. 1–8.

Becker-Asano, Christian and Ipke Wachsmuth (2010). “WASABI as a case study of how misattribution of emotion can be modelled computationally”. In: *A Blueprint for Affective Computing: a Sourcebook and Manual*, pp. 179–193.

Becker-Asano, Christian et al. (2009). “How about laughter? Perceived naturalness of two laughing humanoid robots”. In: *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. IEEE, pp. 1–6.

Becker-Asano, Christian et al. (2010). “Exploring the uncanny valley with Geminoid HI-1 in a real-world application”. In: *Proceedings of IADIS International conference interfaces and human computer interaction*, pp. 121–128.

Bernin, Arne (2011). “Einsatz von 3D-Kameras zur Interpretation von räumlichen Gesten im Smart Home Kontext”. MA thesis. University of Applied Science Hamburg.

Bernin, Arne (2012). “A framework concept for emotion enriched interfaces”. In: *Entertainment Computing-ICEC 2012*, pp. 482–485.

Bernin, Arne et al. (2017). “Towards More Robust Automatic Facial Expression Recognition in Smart Environments”. In: *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, pp. 37–44.

Biddle, Stuart J. H., Kenneth R. Fox, and Stephen H. Boutcher (2000). *Chapter 4: Emotion, mood and physical activity. In: Physical Activity and Psychological Well-being*. Routledge, Psychology Press, p. 63.

Bielenberg, Kai (2016). “Einsatz eines Eyetracker basierten Miningverfahrens für ein Companionsystem”. MA thesis. Hochschule für Angewandte Wissenschaften (HAW) Hamburg.

Bilakhia, Sanjay, Stavros Petridis, and Maja Pantic (2013). “Audiovisual detection of behavioural mimicry”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 123–128.

Borg, Gunnar (2004). “Anstrengungsempfinden und körperliche Aktivität”. In: *Deutsches Ärzteblatt* 101.15, A1016–A1021.

Boucsein, Wolfram (2012). *Electrodermal activity*. Springer Science & Business Media.

Bradley, Margaret M and Peter J Lang (1994). “Measuring emotion: the self-assessment manikin and the semantic differential”. In: *Journal of behavior therapy and experimental psychiatry* 25.1, pp. 49–59.

Brandstätter, Veronika et al. (2013). *Motivation und Emotion: Allgemeine Psychologie für Bachelor*. Springer-Verlag.

Brauer, Henrik Siebo Peter (2014). “Camera based Human Localization and Recognition in Smart Environments”. PhD thesis. dissertation, University of the West of

Scotland, 2014.[Online]. Available: http://users. informatik. haw-hamburg. de/ ubi-comp/arbeiten/phd/brauer. pdf.

Brick, Timothy R, Michael D Hunter, and Jeffrey F Cohn (2009). "Get The FACS Fast: Automated FACS face analysis benefits from the addition of velocity". In: *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. IEEE, pp. 1–7.

Brouwer, Anne-Marie et al. (2013). "Neuroticism, extraversion and stress: Physiological correlates". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 429–434.

Calvo, Rafael A and Sidney D'Mello (2010). "Affect detection: An interdisciplinary review of models, methods, and their applications". In: *IEEE Transactions on affective computing* 1.1, pp. 18–37.

Cardona, J. E. Munoz et al. (2016). "Modulation of Physiological Responses and Activity Levels during Exergame Experiences". In: *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pp. 1–8. DOI: 10.1109/VS-GAMES.2016.7590353.

Chatham, Alan and Florian'Floyd' Mueller (2013). "Adding an interactive display to a public basketball hoop can motivate players and foster community". In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, pp. 667–676.

Chatzakou, Despoina et al. (2013). "Micro-blogging content analysis via emotionally-driven clustering". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 375–380.

Chellali, Ryad and Shannon Hennig (2013). “Is It Time to Rethink Motion Artifacts? Temporal Relationships between Electrodermal Activity and Body Movements in Real-Life Conditions”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 330–335.

Cheng, Jing and Guangyuan Liu (2013). “Computing nonlinear features of skin conductance to build the affective detection model”. In: *Communications, Circuits and Systems (ICCCAS), 2013 International Conference on*. Vol. 2. IEEE, pp. 331–334.

Chittaro, Luca and Riccardo Sioni (2013). “Exploring eye-blink startle response as a physiological measure for affective computing”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 227–232.

Christy, Thomas and Ludmila I Kuncheva (2014). “Technological advancements in affective gaming: A historical survey”. In: *GSTF Journal on Computing (JoC)* 3.4, p. 32.

Constantine, Layale and Hazem Hajj (2012). “A survey of ground-truth in emotion data annotation”. In: *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2012 IEEE International Conference on*. IEEE, pp. 697–702.

Darwin, Charles (1872). *The expression of the emotions in man and animals*. John Murray.

Daum Electronic GmbH (2017). *Kommunikationsprotokoll*. Website. Interface Description. URL: http://www.daum-electronic.de/de/download/bedaprem/SP-KOMM-PM1.pdf.

Dennis, Matt, Judith Masthoff, and Chris Mellish (2013). “Does learner conscientiousness matter when generating emotional support in feedback?” In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 209–214.

Dhall, Abhinav (2013). “Context based facial expression analysis in the wild”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 636–641.

Dickerson, Sally S and Margaret E Kemeny (2004). “Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research.” In: *Psychological bulletin* 130.3, p. 355.

Donato, Gianluca et al. (1999). “Classifying facial actions”. In: *IEEE transactions on pattern analysis and machine intelligence* 21.10, pp. 974–989.

Dorner, D. and K. Hille (1995). “Artificial souls: motivated emotional robots”. In: *1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century*. Vol. 4, 3828–3832 vol.4. DOI: 10.1109/ICSMC.1995.538385.

Dumas, Bruno, Denis Lalanne, and Sharon Oviatt (2009). “Multimodal interfaces: A survey of principles, models and frameworks”. In: *Human machine interaction*. Springer, pp. 3–26.

Ekman, Paul (2004). “Gefühle lesen”. In: *Wie Sie Emotionen erkennen und richtig interpretieren* 2.

– (2016). *Paul Ekman Group LLC*. Website. Last Access 04.10.2016. URL: http://www.paulekman.com/.

Ekman, Paul and Wallace V Friesen (1976). “Measuring facial movement”. In: *Environmental psychology and nonverbal behavior* 1.1, pp. 56–75.

Ekman, Paul, Robert W Levenson, and Wallace V Friesen (1983). “Autonomic nervous system activity distinguishes among emotions”. In: *Science* 221.4616, pp. 1208–1210.

Ellenberg, Jens et al. (2011). “An environment for context-aware applications in smart homes”. In: *International Conference on Indoor Positioning and Indoor Navigation (IPIN), Guimaraes, Portugal*.

Ernst, Andreas and Christian Kublbeck (2011). “Fast face detection and species classification of African great apes”. In: *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*. IEEE, pp. 279–284.

Ernst, Andreas, Tobias Ruf, and Christian Kueblbeck (2009). “A modular framework to detect and analyze faces for audience measurement systems”. In: *2nd Workshop on Pervasive Advertising at Informatik*. Citeseer, pp. 75–87.

Feldman, Pamela J et al. (1999). “Negative emotions and acute physiological responses to stress”. In: *Annals of Behavioral Medicine* 21.3, pp. 216–222.

Friedrichs, Thomas et al. (2015). “Simple Games–Complex Emotions: Automated Affect Detection Using Physiological Signals”. In: *International Conference on Entertainment Computing*. Springer, pp. 375–382.

Garbas, Jens-Uwe et al. (2013). “Towards robust real-time valence recognition from facial expressions for market research applications”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 570–575.

Gilleade, Kiel, Alan Dix, and Jen Allanson (2005). “Affective videogames and modes of affective gaming: assist me, challenge me, emote me”. In: *DiGRA 2005: Changing Views – Worlds in Play.*

Gilleade, Kiel M and Alan Dix (2004). “Using frustration in the design of adaptive videogames”. In: *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology*. ACM, pp. 228–232.

Giraud, Tom et al. (2013a). “Assessing postural control for affect recognition using video and force plates”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 109–115.

Giraud, Tom et al. (2013b). “Multimodal Expressions of Stress during a Public Speaking Task: Collection, Annotation and Global Analyses”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 417–422.

Glowinski, Donald et al. (2013). “Expressive non-verbal interaction in string quartet”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 233–238.

Gosling, Samuel D, Peter J Rentfrow, and William B Swann (2003). “A very brief measure of the Big-Five personality domains”. In: *Journal of Research in personality* 37.6, pp. 504–528.

Grafsgaard, Joseph F et al. (2013). “Automatically recognizing facial indicators of frustration: a learning-centric analysis”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 159–165.

Griffin, Harry J et al. (2013). “Laughter type recognition from whole body motion”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 349–355.

Group, Geneva Emotion Research et al. (2002). “Geneva Appraisal Questionnaire (GAQ): Format, development, and utilization”. In: *Staženo* 16.1, p. 2015.

Gunes, Hatice et al. (2011a). “Emotion representation, analysis and synthesis in continuous space: A survey”. In: *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, p. 827.

– (2011b). “Emotion representation, analysis and synthesis in continuous space: A survey”. In: *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, pp. 827–834.

Hakim, Ayesha, Stephen Marsland, and Hans W Guesgen (2013). “Computational analysis of emotion dynamics”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 185–190.

Hamilton, Iain et al. (2012). “Walk2Build: a GPS game for mobile exergaming with city visualization”. In: *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services companion*. ACM, pp. 17–22.

He, Shan et al. (2013). “Facial expression recognition using deep Boltzmann machine from thermal infrared images”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 239–244.

Healey, Jennifer A and Rosalind W Picard (2005). “Detecting stress during real-world driving tasks using physiological sensors”. In: *IEEE Transactions on intelligent transportation systems* 6.2, pp. 156–166.

Hoda, Mohamad, Rana Alattas, and Abdulmotaleb El Saddik (2013). “Evaluating player experience in cycling exergames”. In: *Multimedia (ISM), 2013 IEEE International Symposium on*. IEEE, pp. 415–420.

Hong, Jin-Hyuk, Julian Ramos, and Anind K Dey (2012). “Understanding physiological responses to stressors during physical activity”. In: *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, pp. 270–279.

Hook, K. (2008). “Knowing, Communication and Experiencing through Body and Emotion”. In: *IEEE Transactions on Learning Technologies* 1.4, pp. 248–259. ISSN: 1939-1382. DOI: 10.1109/TLT.2009.3.

Hoque, Mohammed and Rosalind W Picard (2011). “Acted vs. natural frustration and delight: Many people smile in natural frustration”. In: *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, pp. 354–359.

Hoque, Mohammed E, Rana El Kaliouby, and Rosalind W Picard (2009). “When human coders (and machines) disagree on the meaning of facial affect in spontaneous videos”. In: *International Workshop on Intelligent Virtual Agents*. Springer, pp. 337–343.

Hornschuh, Jonas (2015). “Weiterentwicklung eines Fahrradergometers als intuitive Steuerung für virtuelle Welten”. Bachelor thesis. HAW Hamburg. URL: https://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/bachelor/hornschuh.pdf.

Islam, Monira et al. (2014). “Channel selection and feature extraction for cognitive state estimation with the variation of brain signal”. In: *Electrical Information and Communication Technology (EICT), 2013 International Conference on*. IEEE, pp. 1–6.

James, William (1884). “What is an emotion?” In: *Mind* 34, pp. 188–205.

Jang, E-H et al. (2011). “Identification of the optimal emotion recognition algorithm using physiological signals”. In: *Engineering and Industries (ICEI), 2011 International Conference on*. IEEE, pp. 1–6.

John, Oliver P and Sanjay Srivastava (1999). “The Big Five trait taxonomy: History, measurement, and theoretical perspectives”. In: *Handbook of personality: Theory and research* 2.1999, pp. 102–138.

Jürgen, Bortz and Nicola Döring (2006). *Forschungsmethoden und Evaluation für Human und Sozialwissenschaftler, 4. überarb. Aufl.* Heidelberg: Springer.

Kächele, Markus et al. (2015). “Fusion Mappings for Multimodal Affect Recognition”. In: *Computational Intelligence, 2015 IEEE Symposium Series on*. IEEE, pp. 307–313.

Kaiser, Susanne and T Wehrle (1996). “Situated emotional problem solving in interactive computer games”. In: *In N., H. Frijda (Ed.), Proceedings of the VIXth Conference of the International Society for Research on Emotions*, pp. 276–280.

Kanade, Takeo, Jeffrey F Cohn, and Yingli Tian (2000). “Comprehensive database for facial expression analysis”. In: *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*. IEEE, pp. 46–53.

Kennington, Casey, Spyridon Kousidis, and David Schlangen (2014a). “Multimodal dialogue systems with inprotks and venice”. In: *Proceedings of the 18th SemDial Workshop on the Semantics and Pragmatics of Dialogue (DialWatt). Posters*.

Kennington, Casey, Spyros Kousidis, and David Schlangen (2014b). “InproTKs: A toolkit for incremental situated processing”. In: *Proceedings of SIGdial 2014: Short Papers*.

Kindness, Peter (2013). “Towards a virtual teammate whose support can help alleviate stress in the prehospital care domain”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 683–688.

Kindness, Peter, Chris Mellish, and Judith Masthoff (2013a). “How virtual teammate support types affect stress”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 300–305.

– (2013b). “Identifying and measuring stressors present in pre-hospital care”. In: *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), pp. 270–273.

Kipp, Andreas and Franz Kummert (2014). “Dynamic dialog system for human robot collaboration: playing a game of pairs”. In: *Proceedings of the second international conference on Human-agent interaction*. ACM, pp. 225–228.

Kipp, M (2001). “Anvil-a generic annotation tool for multimodal dialogue" in procs of 7th European Conference on Speech Communication and Technology”. In:

Kirschbaum, Clemens, K-M Pirke, and Dirk H Hellhammer (1993). “The Trier Social Stress Test – a tool for investigating psychobiological stress responses in a laboratory setting”. In: *Neuropsychobiology* 28.1-2, pp. 76–81.

Kleinginna Jr, Paul R and Anne M Kleinginna (1981). “A categorized list of emotion definitions, with suggestions for a consensual definition”. In: *Motivation and emotion* 5.4, pp. 345–379.

Kletz, Florian (2016). “Thermalbilder im Kontext von Companion Systemen”. In: URL: https://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2016-proj/kletz2.pdf.

Kletz, Florian and Jorin Kleimann (2016). “Evaluierung von Lichtfeld- und Thermogra?ekameras im Human-computer interaction Kontext”. In: URL: https://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2016-proj/kletz-kleimann.pdf.

Koelstra, Sander et al. (2012). “Deap: A database for emotion analysis; using physiological signals”. In: *IEEE Transactions on Affective Computing* 3.1, pp. 18–31.

Kousidis, Spyros et al. (2014). “A Multimodal In-Car Dialogue System That Tracks The Driver’s Attention”. In: *Proceedings of the 16th International Conference on Multimodal Interaction*. ACM, pp. 26–33.

Krieger, Vincent, Elise Lallart, and Roland Jouvent (2013). “Bodily Manifestations of Affects: The Example of Gait and Virtual Reality”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 179–184.

Küblbeck, Christian and Andreas Ernst (2006). “Face detection and tracking in video sequences using the modifiedcensus transformation”. In: *Image and Vision Computing* 24.6, pp. 564–572.

Kuckartz, Udo (2014). *Mixed Methods: Methodologie, Forschungsdesigns und Analyseverfahren*. Springer.

Kudielka, Brigitte M et al. (2007). "Ten years of research with the Trier Social Stress Test-revisited". In: *Social neuroscience: Integrating biological and psychological explanations of social behavior*, pp. 56–83.

Lang, Peter J, Margaret M Bradley, and Bruce N Cuthbert (2005). "International affective picture system (IAPS): Affective ratings of pictures and instruction manual". In: *Technical report A-6*.

Lange, Carl (1887, Danish org. 1885). *Üeber Gemütsbewegungen*. Ed. by Theodor Thomas. Leipzig.

Lathia, Neal et al. (2016). "Happier People Live More Active Lives: Using smartphones to link happiness and physical activity". In:

Lazarus, R.S. (1999). *Stress and Emotion: A New Synthesis*. Springer Publishing Company. ISBN: 9780826112507. URL: https://books.google.ca/books?id=5pb-RAAACAAJ.

Lee, EM et al. (2010). "Glass-type wireless PPG measuring system". In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. IEEE, pp. 1433–1436.

Lee, Jong-Seok and Cheol Hoon Park (2008). "Robust audio-visual speech recognition based on late integration". In: *IEEE Transactions on Multimedia* 10.5, pp. 767–779.

Li, Longfei et al. (2013). "Hybrid Deep Neural Network–Hidden Markov Model (DNN-HMM) Based Speech Emotion Recognition". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 312–317.

Liebold, Benny and Peter Ohler (2013). “Multimodal Emotion Expressions of Virtual Agents, Mimic and Vocal Emotion Expressions and Their Effects on Emotion Recognition”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 405–410.

Littlewort, Gwen et al. (2006). “Dynamics of facial expression extracted automatically from video”. In: *Image and Vision Computing* 24.6, pp. 615–625.

Littlewort, Gwen et al. (2011). “The computer expression recognition toolbox (CERT)”. In: *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, pp. 298–305.

Liu, He, Yadong Wang, and Lei Wang (2013). “FPGA-based remote pulse rate detection using photoplethysmographic imaging”. In: *2013 IEEE International Conference on Body Sensor Networks*. IEEE, pp. 1–5.

Lütkebohle, Ingo et al. (2010). “The Bielefeld anthropomorphic robot head "Flobi"”. In: *Robotics and automation (ICRA), 2010 IEEE international conference on*. IEEE, pp. 3384–3391.

Malaka, Rainer (2014). “How computer games can improve your health and fitness”. In: *International Conference on Serious Games*. Springer, pp. 1–7.

Matthiessen, Erik (2015). “Entwicklung einer Gangschaltungs- und Bremsensteuerung für ein Fahrradergometer”. In: bachelor thesis. URL: https://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/bachelor/matthiessen.pdf.

Mazzei, Daniele et al. (2012). “Robotic social therapy on children with autism: preliminary evaluation through multi-parametric analysis”. In: *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. IEEE, pp. 766–771.

McCrae, Robert R (1990). "Controlling neuroticism in the measurement of stress". In: *Stress Medicine* 6.3, p. 237.

McDuff, Daniel et al. (2013). "Measuring voter's candidate preference based on affective responses to election debates". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 369–374.

McKeown, Gary et al. (2012). "The semaine database: Annotated multimodal records of emotionally colored conversations between a person and a limited agent". In: *IEEE Transactions on Affective Computing* 3.1, pp. 5–17.

McKeown, Gary et al. (2013). "Human perception of laughter from context-free whole body motion dynamic stimuli". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 306–311.

Mey, Günter and Katja Mruck (2010). *Handbuch qualitative Forschung in der Psychologie*. Vol. 1. Springer.

Mori, Masahiro, Karl F MacDorman, and Norri Kageki (2012, Japanese orig. 1970). "The uncanny valley [from the field]". In: *IEEE Robotics & Automation Magazine* 19.2, pp. 98–100.

Müller, Larissa (2013). "Emotionale Modellierung in Mensch-Maschine-Interaktionen". MA thesis. University of Applied Sciences Hambuirg (HAW). URL: https://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/master/mueller.pdf.

Müller, Larissa et al. (2012a). "Emotion Sensitive Active Surfaces". In: Workshop Article of the Track Emotion and Computing: presented at the 35th German Conference on Artificial Intelligence (http://www.dfki.de/KI2012/. URL: http://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/papers/KI2012.pdf.

Müller, Larissa et al. (2012b). "Emotional interaction with surfaces-works of design and computing". In: *International Conference on Entertainment Computing*. Springer, pp. 457–460.

Müller, Larissa et al. (2015). "EmotionBike: a study of provoking emotions in cycling exergames". In: *International Conference on Entertainment Computing*. Springer, pp. 155–168.

Müller, Larissa et al. (2016). "Physiological data analysis for an emotional provoking exergame". In: *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on*. IEEE, pp. 1–8.

Müller, Larissa et al. (2017a). "Emotional Journey for an Emotion Provoking Cycling Exergame". In: *Soft Computing & Machine Intelligence (ISCMI 2017), 2017 IEEE 4th Intl. Conference on*. IEEE, pp. 104–108.

– (2017b). "Enhancing Exercise Experience with Individual Multi-Emotion Provoking Game Elements". In: *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on*. IEEE, pp. 1–8.

Munia, Tamanna Tabassum Khan et al. (2012). "Mental states estimation with the variation of physiological signals". In: *Informatics, Electronics & Vision (ICIEV), 2012 International Conference on*. IEEE, pp. 800–805.

Nacke, Lennart Erik et al. (2011). "Biofeedback game design: using direct and indirect physiological control to enhance game interaction". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, pp. 103–112.

Nakajima, K, T Tamura, and H Miike (1996). "Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique". In: *Medical engineering & physics* 18.5, pp. 365–372.

Nasoz, Fatma and Mehmet Bayburt (2009). “Affectively intelligent user interfaces for enhanced e-learning applications”. In: *Human Centered Design*, pp. 765–774.

Nasoz, Fatma, Christine L Lisetti, and Athanasios V Vasilakos (2010). “Affectively intelligent and adaptive car interfaces”. In: *Information Sciences* 180.20, pp. 3817–3836.

Negini, Faham, Regan L Mandryk, and Kevin G Stanley (2014). “Using affective state to adapt characters, NPCs, and the environment in a first-person shooter game”. In: *Games Media Entertainment (GEM), 2014 IEEE*. IEEE, pp. 1–8.

Nogueira, Pedro A. et al. (2016). “Vanishing scares: biofeedback modulation of affective player experiences in a procedural horror game”. In: *Journal on Multimodal User Interfaces* 10.1, pp. 31–62. ISSN: 1783-8738. DOI: 10.1007/s12193-015-0208-1. URL: https://doi.org/10.1007/s12193-015-0208-1.

O’Brien, Heather L and Elaine G Toms (2010). “The development and evaluation of a survey to measure user engagement”. In: *Journal of the American Society for Information Science and Technology* 61.1, pp. 50–69.

Ochs, Magalie, Ken Prepin, and Catherine Pelachaud (2013). “From emotions to interpersonal stances: Multi-level analysis of smiling virtual characters”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 258–263.

Organization, World Health (2001). *The World Health Report 2001: Mental health: new understanding, new hope*. World Health Organization.

Ortony, Andrew, Gerald L Clore, and Allan Collins (1988). *The Congnitive Structure of Emotions*. Cambridge University Press. URL: https://books.google.de/books?id=dA3JEEAp6TsC&printsec=frontcover&hl=de#v=onepage&q&f=false.

Otto, Kjell and Sören Voskuhl (2010). "Entwicklung einer Architektur für den Living Place Hamburg". In: University of Applied Sciences Hamburg. URL: https://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2010-proj1/otto_voskuhl.pdf.

Pantic, Maja et al. (2005). "Web-based database for facial expression analysis". In: *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*. IEEE, 5–pp.

Parnandi, Avinash, Youngpyo Son, and Ricardo Gutierrez-Osuna (2013). "A control-theoretic approach to adaptive physiological games". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 7–12.

Peter, Christian and Russell Beale (2008). *Affect and Emotion in Human-Computer Interaction: From Theory to Applications*. 1st. Springer Publishing Company, Incorporated. ISBN: 3540850988, 9783540850984.

Petridis, Stavros, Maelle Leveque, and Maja Pantic (2013). "Audiovisual detection of laughter in human-machine interaction". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 129–134.

Petridis, Stavros, Brais Martinez, and Maja Pantic (2013). "The MAHNOB laughter database". In: *Image and Vision Computing* 31.2, pp. 186–202.

Pfister, Tomas and Peter Robinson (2011). "Real-time recognition of affective states from nonverbal features of speech and its application for public speaking skill analysis". In: *IEEE Transactions on Affective Computing* 2.2, pp. 66–78.

Picard, Rosalind W. (1997a). *Affective Computing*. Cambridge, MA, USA: MIT Press. ISBN: 0-262-16170-2.

– (1997b). *Affective Computing*. Cambridge, MA, USA: MIT Press, xi, Preface. ISBN: 0-262-16170-2.

Picard, Rosalind W and Shaundra Bryant Daily (2005). “Evaluating affective interactions: Alternatives to asking what users feel”. In: *CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches*. ACM New York, NY, pp. 2119–2122.

Plarre, Kurt et al. (2011). “Continuous inference of psychological stress from sensory measurements collected in the natural environment”. In: *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*. IEEE, pp. 97–108.

Plutchik, Robert (2001). “The Nature of Emotions”. In: *American Scientist* 89.4, pp. 344–350.

Prendinger, Helmut, Christian Becker, and Mitsuru Ishizuka (2006). “A STUDY IN USERS’PHYSIOLOGICAL RESPONSE TO AN EMPATHIC INTERFACE AGENT”. In: *International Journal of Humanoid Robotics* 3.03, pp. 371–391.

Pütten, Astrid M von der et al. (2011). “An android in the field”. In: *Proceedings of the 6th international conference on Human-robot interaction*. ACM, pp. 283–284.

Raaijmakers, SF et al. (2013). “Heart rate variability and skin conductance biofeedback: A triple-blind randomized controlled study”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 289–293.

Rabiner, Lawrence R and Bernard Gold (1975). “Theory and application of digital signal processing”. In: *Englewood Cliffs, NJ, Prentice-Hall, Inc., 1975. 777 p.* 1.

Reisenzein, Rainer (2006). “Arnold’s theory of emotion in historical perspective”. In: *Cognition and Emotion* 20.7, pp. 920–951.

Russel, James A. (1980). “A Circumplex Model of Affect”. In: *Journal of Personality and Social Psychology* 39.6, pp. 1161–1178.

Russell, James A and Albert Mehrabian (1977). “Evidence for a three-factor theory of emotions”. In: *Journal of research in Personality* 11.3, pp. 273–294.

Russell, James A and Kaori Sato (1995). “Comparing emotion words between languages”. In: *Journal of Cross-Cultural Psychology* 26.4, pp. 384–391.

Samadani, Ali-Akbar et al. (2013). “Laban effort and shape analysis of affective hand and arm movements”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 343–348.

Sapru, Ashtosh and Hervé Bourlard (2013). “Investigating the impact of language style and vocal expression on social roles of participants in professional meetings”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 324–329.

Sareen, Meghna et al. (2008). “Nadi Yantra: a robust system design to capture the signals from the radial artery for non-invasive diagnosis”. In: *2008 2nd International Conference on Bioinformatics and Biomedical Engineering*. IEEE, pp. 1387–1390.

Sarris, Viktor and Siegbert Reiß (2005). *Kurzer Leitfaden der Experimentalpsychologie*. Pearson Studium München.

Satow, L (2012). “Big-Five-Persönlichkeitstest(B5T): Test-und Skalendokumentation”. In: Online im Internet: URL: http://www.drsatow.de.

Sauter, Disa A et al. (2010). “Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations”. In: *Proceedings of the National Academy of Sciences* 107.6, pp. 2408–2412.

Sawka, Michael N et al. (2007). “American College of Sports Medicine position stand. Exercise and fluid replacement.” In: *Medicine and science in sports and exercise* 39.2, pp. 377–390.

Schaaff, Kristina and Marc TP Adam (2013). “Measuring emotional arousal for online applications: evaluation of ultra-short term heart rate variability measures”. In: *Affective Com-*

puting and Intelligent Interaction (ACII), 2013 Humaine Association Conference on. IEEE, pp. 362–368.

Scherer, Klaus R. (2000). “Emotion”. In: *M. Hewstone & W. Stroebe (Eds.). Introduction to Social Psychology: A European perspective*. 3rd ed. Oxford: Blackwell, pp. 151–191.

Scherer, Klaus R (2005). “What are emotions? And how can they be measured?” In: *Social science information* 44.4, pp. 695–729.

Schneider, Tamera R et al. (2012). “The influence of neuroticism, extraversion and openness on stress responses”. In: *Stress and Health* 28.2, pp. 102–110.

Sharma, Nandita and Tom Gedeon (2013). “Modeling stress recognition in typical virtual environments”. In: *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. IEEE, pp. 17–24.

Sharma, Nandita et al. (2013). “Modeling stress using thermal facial patterns: A spatio-temporal approach”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 387–392.

Shih, Sung-Tsun, Chian-Yi Chao, and Chin-Ming Hsu (2011). “RFID based physical fitness condition measurement system”. In: *Networked Computing and Advanced Information Management (NCM), 2011 7th International Conference on*. IEEE, pp. 284–288.

Sidney, K Dmello et al. (2005). “Integrating affect sensors in an intelligent tutoring system”. In: *Affective Interactions: The Computer in the Affective Loop Workshop at*, pp. 7–13.

Singh, Hari, Rajesh Singla, and R Jha (2009). “The Effect of Mental States on Blood Pressure and Electrocardiogram”. In: *2009 3rd International Conference on Bioinformatics and Biomedical Engineering*. IEEE, pp. 1–4.

Smith, Steven W. (1997). “The Scientist & Engineer’s Guide to Digital Signal Processing”. In: ed. by California Technical Pub, pp. 274 –283.

Snel, John and Charlie Cullen (2013). “Judging emotion from low-pass filtered naturalistic emotional speech”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 336–342.

Soleimani, Ahmad and Ziad Kobti (2013). “Event-driven fuzzy automata for tracking changes in the emotional behavior of affective agents”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 221–226.

Steunebrink, Bas R, Mehdi Dastani, and John-Jules Ch Meyer (2009). “The OCC model revisited”. In: *Proc. of the 4th Workshop on Emotion and Computing*. Vol. 65, pp. 2047–2056.

Strack, Fritz, Leonard L Martin, and Sabine Stepper (1988). “Inhibiting and facilitating conditions of the human smile: a nonobtrusive test of the facial feedback hypothesis.” In: *Journal of personality and social psychology* 54.5, pp. 768–777.

Stratou, Giota et al. (2013). “Automatic nonverbal behavior indicators of depression and PTSD: Exploring gender differences”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 147–152.

Sun, Xiaofan et al. (2011). “A multimodal database for mimicry analysis”. In: *Affective Computing and Intelligent Interaction*, pp. 367–376.

Süssenbach, Luise et al. (2014). “A robot as fitness companion: towards an interactive action-based motivation model”. In: *Robot and Human Interactive Communication, 2014 RO-MAN: The 23rd IEEE International Symposium on*. IEEE, pp. 286–293.

Tao, Jianhua and Tieniu Tan (2005). “Affective Computing: A Review”. In: *Affective Computing and Intelligent Interaction: First International Conference, ACII 2005, Beijing, China, October 22-24, 2005. Proceedings*. Ed. by Jianhua Tao, Tieniu Tan, and Rosalind W. Picard. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 981–995. ISBN: 978-3-540-

32273-3. DOI: 10.1007/11573548_125. URL: https://doi.org/10.1007/11573548_125.

Thayer, R.E. (1990). *The Biopsychology of Mood and Arousal*. Oxford University Press. ISBN: 9780195361759. URL: https://books.google.de/books?id=ORiwiDNqcbEC.

The American Psychological Association. http://goo.gl/S2Jgm. Accessed: 2016-09-05.

Togelius, Julian and Georgios N Yannakakis (2016). "Emotion-Driven Level Generation". In: *Emotion in Games*. Springer, pp. 155–166.

Tsalamlal, Mohamed Yassine et al. (2013). "EmotionAir: perception of emotions from air jet based tactile stimulation". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 215–220.

Urbain, Jérôme, Hüseyin Çakmak, and Thierry Dutoit (2013). "Automatic phonetic transcription of laughter and its application to laughter synthesis". In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 153–158.

Vachiratamporn, Vanus et al. (2014). "An implementation of affective adaptation in survival horror games". In: *Computational Intelligence and Games (CIG), 2014 IEEE Conference on*. IEEE, pp. 1–8.

Valstar, Michel and Maja Pantic (2010). "Induced disgust, happiness and surprise: an addition to the mmi facial expression database". In: *Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect*, p. 65.

Velloso, Eduardo, Andreas Bulling, and Hans Gellersen (2013). "AutoBAP: Automatic coding of body action and posture units from wearable sensors". In: *Affective Computing and*

Intelligent Interaction (ACII), 2013 Humaine Association Conference on. IEEE, pp. 135–140.

Wahlster, W. and A. Kobsa (1986). "Dialogue-based user models". In: *Proceedings of the IEEE* 74.7, pp. 948–960. ISSN: 0018-9219. DOI: 10.1109/PROC.1986.13574.

Wahlster, Wolfgang (2006). "Dialogue systems go multimodal: the SmartKom experience". In: *SmartKom: foundations of multimodal dialogue systems*. Springer, pp. 3–27.

Wallbott, Harald G and Klaus R Scherer (1991). "Stress specificities: Differential effects of coping style, gender, and type of stressor on autonomic arousal, facial expression, and subjective feeling." In: *Journal of Personality and Social Psychology* 61.1, p. 147.

Walmink, Wouter, Danielle Wilde, and Florian'Floyd' Mueller (2014). "Displaying heart rate data on a bicycle helmet to support social exertion experiences". In: *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*. ACM, pp. 97–104.

Wang, Ning and Stacy Marsella (2006). "Introducing EVG: An emotion evoking game". In: *International Workshop on Intelligent Virtual Agents*. Springer, pp. 282–291.

Warburton, Darren ER et al. (2007). "The health benefits of interactive video game exercise". In: *Applied Physiology, Nutrition, and Metabolism* 32.4, pp. 655–663.

Watson, David, Lee A Clark, and Auke Tellegen (1988). "Development and validation of brief measures of positive and negative affect: the PANAS scales." In: *Journal of personality and social psychology* 54.6, p. 1063.

Weiser, Mark (1991). "The Computer for the 21 st Century". In: *Scientific american* 265.3, pp. 94–105.

Westerink, J., M. Krans, and M. Ouwerkerk (2011). *Sensing Emotions: The impact of context on experience measurements*. Philips Research Book Series. Springer Netherlands. ISBN: 9789048132584.

Wundt, Wilhelm (1863). “Vorlesungen über die Menschen-und Thierseele”. In: *Leipzig: Voss*.

Yamauchi, Takashi (2013). “Mouse trajectories and state anxiety: Feature selection with random forest”. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*. IEEE, pp. 399–404.

Yannakakis, Georgios N, Hector P Martinez, and Maurizio Garbarino (2016). “Psychophysiology in games”. In: *Emotion in Games*. Springer, pp. 119–137.

Yu, Changrong, Jiehan Zhou, and Jukka Riekki (2009). “Expression and analysis of emotions: Survey and experiment”. In: *Ubiquitous, Autonomic and Trusted Computing, 2009. UIC-ATC’09. Symposia and Workshops on*. IEEE, pp. 428–433.

Yu, M et al. (2004). “Sensory-motor coordination using visual stimulation”. In: *Industrial Electronics Society, 2004. IECON 2004. 30th Annual Conference of IEEE*. Vol. 2. IEEE, pp. 1070–1075.

Zagaria, Sebastain (2017). “Emotional adäquat reagierende KI-Agneten zur Erhöhung von Immersion in Computerspielen”. MA thesis. Hochschule für Angewandte Wissenschaften (HAW) Hamburg. URL: https://users.informatik.haw-hamburg.de/~ubicomp/arbeiten/master/zagaria.pdf.

Zalis, E.G. and M.B. Conover (1972). *Understanding Electrocardiography: Physiological and Interpretive Concepts*. Mosby. ISBN: 9780801656743. URL: https://books.google.de/books?id=FWC8AAAAIAAJ.

Zeng, Zhihong et al. (2009). “A survey of affect recognition methods: Audio, visual, and spontaneous expressions”. In: *IEEE transactions on pattern analysis and machine intelligence* 31.1, pp. 39–58.

Appendices

A. Questionnaires

At the beginning every participants was asked to fill out questionnaires. In the following pictures of the web interface are presented. The applied questionnaires are displayed in their german original.

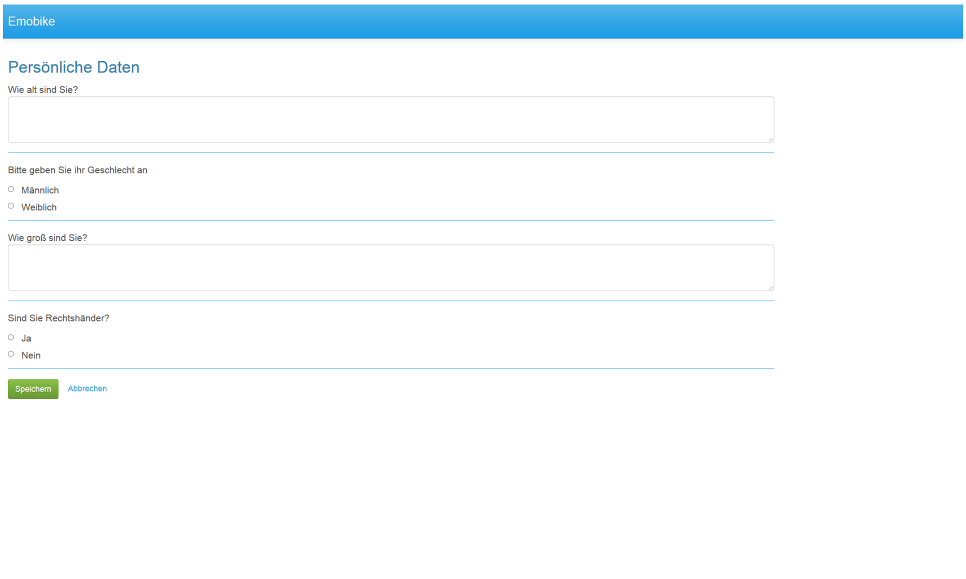
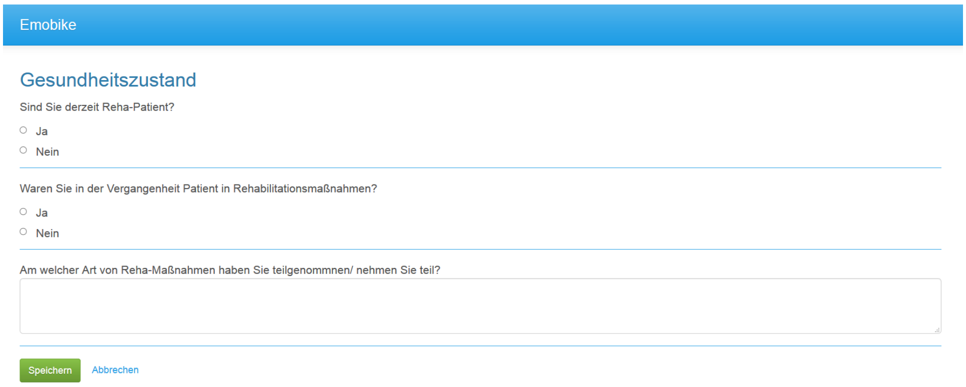


Figure A.1.: Questionnaire Part 1: Personal Information



Emobike

Gesundheitszustand

Sind Sie derzeit Reha-Patient?

- ○ Ja
- ○ Nein

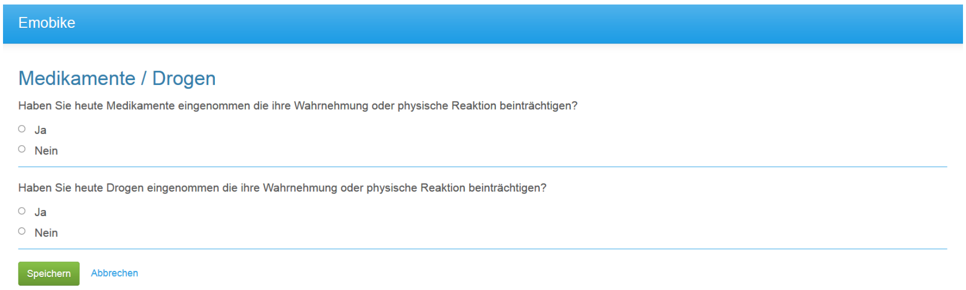
Waren Sie in der Vergangenheit Patient in Rehabilitationsmaßnahmen?

- ○ Ja
- ○ Nein

Am welcher Art von Reha-Maßnahmen haben Sie teilgenommnen/ nehmen Sie teil?

Speichern Abbrechen

Figure A.2.: Questionnaire Part 2: Personal Information



Emobike

Medikamente / Drogen

Haben Sie heute Medikamente eingenommen die ihre Wahrnehmung oder physische Reaktion beinträchtigen?

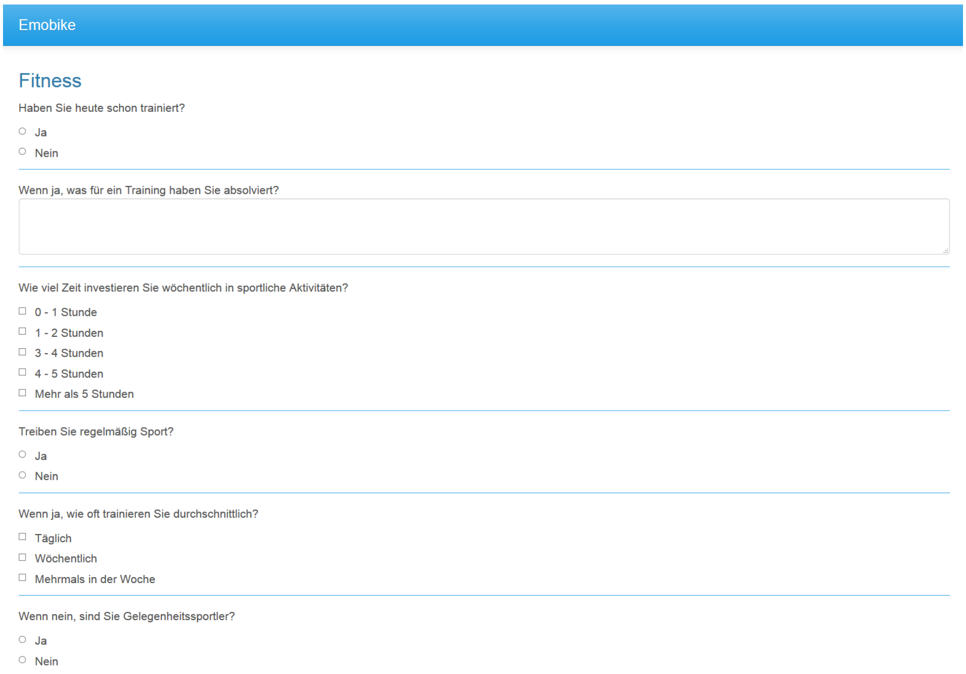
- ○ Ja
- ○ Nein

Haben Sie heute Drogen eingenommen die ihre Wahrnehmung oder physische Reaktion beinträchtigen?

- ○ Ja
- ○ Nein

Speichern Abbrechen

Figure A.3.: Questionnaire Part 3: Personal Information



Emobike

Fitness

Haben Sie heute schon trainiert?

- ○ Ja
- ○ Nein

Wenn ja, was für ein Training haben Sie absolviert?

Wie viel Zeit investieren Sie wöchentlich in sportliche Aktivitäten?

- ☐ 0 - 1 Stunde
- ☐ 1 - 2 Stunden
- ☐ 3 - 4 Stunden
- ☐ 4 - 5 Stunden
- ☐ Mehr als 5 Stunden

Treiben Sie regelmäßig Sport?

- ○ Ja
- ○ Nein

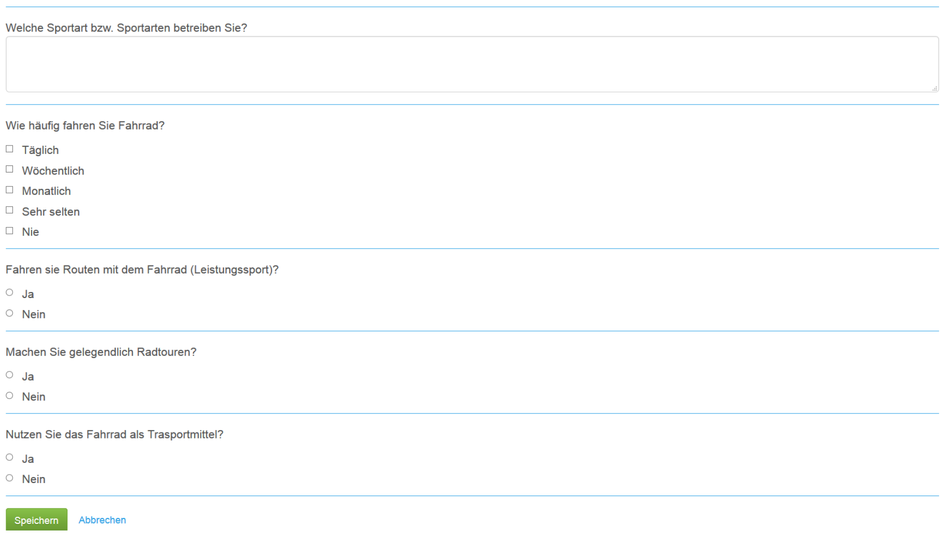
Wenn ja, wie oft trainieren Sie durchschnittlich?

- ☐ Täglich
- ☐ Wöchentlich
- ☐ Mehrmals in der Woche

Wenn nein, sind Sie Gelegenheitssportler?

- ○ Ja
- ○ Nein

Figure A.4.: Questionnaire Part 4: Fitness Condition



Welche Sportart bzw. Sportarten betreiben Sie?

Wie häufig fahren Sie Fahrrad?

- ☐ Täglich
- ☐ Wöchentlich
- ☐ Monatlich
- ☐ Sehr selten
- ☐ Nie

Fahren sie Routen mit dem Fahrrad (Leistungssport)?

- ○ Ja
- ○ Nein

Machen Sie gelegendlich Radtouren?

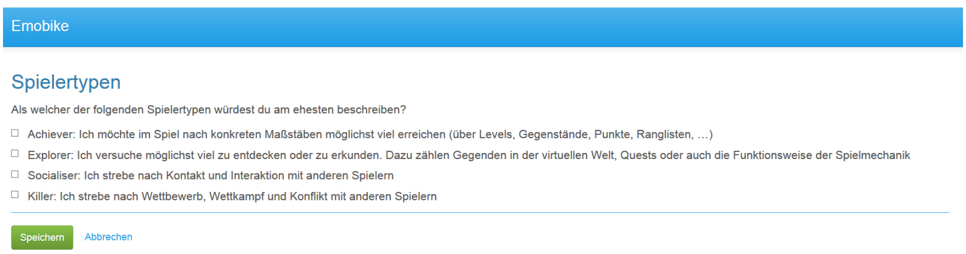
- ○ Ja
- ○ Nein

Nutzen Sie das Fahrrad als Trasportmittel?

- ○ Ja
- ○ Nein

Speichern Abbrechen

Figure A.5.: Questionnaire Part 5: Fitness Condition



Emobike

Spielertypen

Als welcher der folgenden Spielertypen würdest du am ehesten beschreiben?

- ☐ Achiever: Ich möchte im Spiel nach konkreten Maßstäben möglichst viel erreichen (über Levels, Gegenstände, Punkte, Ranglisten, ...)
- ☐ Explorer: Ich versuche möglichst viel zu entdecken oder zu erkunden. Dazu zählen Gegenden in der virtuellen Welt, Quests oder auch die Funktionsweise der Spielmechanik
- ☐ Socialiser: Ich strebe nach Kontakt und Interaktion mit anderen Spielern
- ☐ Killer: Ich strebe nach Wettbewerb, Wettkampf und Konflikt mit anderen Spielern

Speichern Abbrechen

Figure A.6.: Questionnaire Part 6: Game Experience

Emobike

Game Erfahrung

Spielen Sie Videospiele?

- ○ Ja
- ○ Nein

Wie viel Zeit investieren Sie wöchentlich in Videospiele?

- ☐ 0 - 1h
- ☐ 2h - 4h
- ☐ 5h - 6h
- ☐ 7h oder mehr

Welche Art von Videospielen spielen Sie?

- ☐ Action
- ☐ Jump and Run
- ☐ Sport
- ☐ Strategie Spiele
- ☐ Mobile Open World Game
- ☐ Rennspiele
- ☐ Multiplayer
- ☐ Adventure
- ☐ Simulation
- ☐ Gelegenheits Spiele
- ☐ First Person Shooter

Wie gut schätzen Sie ihre Spielfertigkeit ein?

- ☐ 1 Sehr gering
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5 Sehr hoch

Figure A.7.: Questionnaire Part 7: Game Experience

Welche Erfahrungen haben Sie mit neuen Technologien als Game Controller?

- ☐ Augmented Reality
- ☐ Guitar Hero (Gitarre als Controller)
- ☐ Microsoft Kinect (Kamera basierte Controller)
- ☐ Wii Controller
- ☐ Wii Balance-Board
- ☐ Lenkrad (z. B. für Rennspiele)
- ☐ Virtual Reality Brillen (z. B. Oculus Rift)

Speichern Abbrechen

Figure A.8.: Questionnaire Part 8: Gamer Type

B. Self-Assessment

The emotion assessment was translated in relation to Figure 5.1, although the german original might differ in detail for emotional manifestations.

ID	Fork Scene	Fun	Forest Scene	Fun	Did you feel threatened by the Environment?	Saw Scene	Coin Collection	Zombies
1	surprise	3	surprise	1	no	X	confused	surprise
2	happy	4	scared	4	no	happy	happy	scared
3	X	4	scared	4	no	curious	happy	scared
4	confused	3	scared	3	no	X	X	scared
5	happy	2	fear	2	yes	challenge	happy	scared
6	surprise	3	scared	4	no	X	happy	scared
7	X	2	scared	3	no	happy	X	scared
8	challenge	3	scared	1	yes	unsettled	X	scared
9	happy	3	happy	2	no	surprise	happy	scared
10	challenge	4	scared	4	yes	fear	happy	stress
11	challenge	4	stress	5	yes	fear	challenge	scared
12	stress	2	challenge	3	no	surprise	challenge	scared
13	surprise	4	scared	5	no	challenge	challenge	scared
14	bored	1	happy	4	no	surprise	happy	happy
15	confused	2	scared	3	yes	challenge	challenge	scared
16	effort	1	happy	3	yes	X	X	scared
17	surprise	4	scared	4	no	surprise	happy	scared
18	surprise	3	scared	3	no	X	angry	scared
19	confused	3	scared	2	no	happy	happy	scared
20	surprise	3	scared	3	yes	fear	happy	scared
21	confused	4	stress	4	yes	surprise	happy	scared
22	happy	4	surprise	5	no	surprise	happy	scared
23	surprise	4	scared	4	yes	happy	happy	scared
24	surprise	4	scared	4	no	confused	X	scared
25	surprise	3	stress	4	yes	surprise	happy	scared

Table B.1.: Self-Assessment: **Fork Scene** and the **Forest Scene**

ID	Challenge	Fun	Landing	Teleport	Scene End	Falling First	Boost First
1	confused	2	happy	X	happy	confused	happy
2	happy	5	happy	X	sad	X	confused
3	frustration	2	X	happy	happy	surprise	X
4	effort	4	surprise	X	X	confused	X
5	challenge	4	happy	X	happy	challenge	X
6	annoyed	2	X	X	X	frustration	X
7	challenge	3	challenge	X	happy	X	X
8	happy	4	surprise	X	happy	happy	annoyed
9	annoyed	2	happy	X	X	happy	X
10	annoyed	2	X	happy	happy	confused	X
11	effort	3	happy	X	confused	sad	happy
12	frustration	4	frustration	frustration	frustration	frustration	X
13	challenge	5	happy	X	happy	challenge	surprise
14	challenge	2	happy	X	X	X	surprise
15	frustration	2	happy	X	X	frustration	happy
16	frustration	2	X	happy	happy	happy	happy
17	happy	3	X	X	X	challenge	X
18	challenge	4	surprise	X	challenge	X	X
19	frustration	1	surprise	X	happy	challenge	surprise
20	challenge	4	happy	X	X	stress	challenge
21	surprise	3	surprise	X	X	X	X
22	challenge	3	happy	X	happy	challenge	fear
23	challenge	5	happy	X	X	surprise	X
24	challenge	2	X	frustration	frustration	challenge	X
25	frustration	2	happy	X	happy	challenge	X

Table B.2.: Self-Assessment: **Challenge Scene** Part 1

ID	Falling Repetition	Boost Repetition	Did you feel Challenged	Has your Ambition been awakened?	Did you feel Frustrated
1	confused	happy	yes	X	no
2	X	X	yes	yes	no
3	frustration	frustration	yes	yes	yes
4	X	X	yes	yes	no
5	challenge	X	yes	yes	no
6	X	X	no	no	yes
7	X	X	yes	no	yes
8	challenge	X	yes	yes	no
9	annoyed	annoyed	yes	yes	yes
10	annoyed	annoyed	yes	no	yes
11	annoyed	bored	yes	yes	yes
12	annoyed	stress	yes	yes	yes
13	X	X	yes	yes	no
14	frustration	challenge	yes	yes	no
15	angry	X	yes	no	yes
16	frustration	frustration	yes	yes	yes
17	X	X	yes	yes	no
18	X	X	yes	yes	no
19	frustration	angry	yes	yes	yes
20	effort	happy	yes	yes	no
21	X	X	no	yes	no
22	frustration	fear	yes	yes	yes
23	challenge	effort	yes	yes	no
24	frustration	challenge	yes	yes	yes
25	frustration	frustration	yes	yes	yes

Table B.3.: Self-Assessment: **Challenge Scene** Part 2

ID	Teddy	Fun	Saw Teddy	Teddy Hit First	Teddy Hit Repetition
1	sad	1	happy	sad	X
2	annoyed	3	confused	confused	X
3	X	3	happy	surprise	X
4	effort	3	confused	stress	stress
5	scared	2	happy	scared	sad
6	happy	5	happy	happy	happy
7	happy	5	happy	happy	happy
8	happy	3	confused	sad	X
9	bored	1	X	scared	annoyed
10	surprise	4	happy	scared	annoyed
11	frustration	3	fear	disgust	sad
12	stress	2	happy	stress	scared
13	confused	4	surprise	confused	challenge
14	happy	1	happy	happy	bored
15	confused	3	happy	surprise	effort
16	sad	2	surprise	sad	sad
17	X	3	surprise	sad	frustration
18	challenge	5	challenge	X	X
19	happy	5	happy	happy	happy
20	fear	2	happy	sad	sad
21	disgust	2	stress	disgust	X
22	confused	1	surprise	scared	bored
23	confused	3	confused	scared	effort
24	confused	3	fear	sad	sad
25	sad	1	surprise	sad	sad

Table B.4.: Self-Assessment: **Teddy Scene**

ID	Frozensea	Fun	Coin	Schnowmen	Ice	Falling	Scene End
1	challenge	3	happy	surprise	stress	frustration	X
2	happy	2	happy	confused	confused	sad	X
3	challenge	3	happy	scared	challenge	X	happy
4	effort	4	happy	scared	challenge	effort	happy
5	challenge	3	happy	scared	annoyed	X	happy
6	annoyed	3	bored	annoyed	X	X	X
7	concentrated	3	happy	X	challenge	X	happy
8	scared	1	happy	scared	effort	X	happy
9	frustration	3	happy	annoyed	X	frustration	happy
10	annoyed	3	happy	angry	annoyed	annoyed	happy
11	stress	4	happy	scared	angry	annoyed	happy
12	challenge	4	happy	scared	confused	annoyed	happy
13	challenge	4	happy	scared	challenge	X	happy
14	effort	2	happy	X	X	X	happy
15	challenge	4	challenge	surprise	effort	X	happy
16	stress	3	happy	scared	stress	frustration	happy
17	challenge	4	happy	scared	X	surprise	happy
18	stress	4	stress	scared	challenge	X	happy
19	stress	4	happy	fear	frustration	angry	happy
20	challenge	4	happy	X	confused	X	happy
21	challenge	3	happy	stress	scared	sad	happy
22	confused	4	happy	scared	X	annoyed	happy
23	challenge	4	happy	scared	X	X	sad
24	challenge	4	happy	scared	effort	X	happy
25	happy	4	happy	scared	challenge	sad	happy

Table B.5.: Self-Assessment: **Frozensea Scene**

ID	Cliff	Fun	Saw Scene	Falling	Saw Rock	Rock Hit	Bridge	Scene End
1	stress	4	frustration	frustration	stress	confused	effort	happy
2	happy	4	neugierig	X	stress	X	happy	happy
3	ambition	4	happy	frustration	surprise	sad	challenge	happy
4	challenge	4	X	surprise	effort	X	effort	happy
5	annoyed	3	X	annoyed	surprise	annoyed	challenge	happy
6	challenge	3	challenge	annoyed	bored	frustration	challenge	happy
7	challenge	4	Ehrgeiz	angry	X	angry	challenge	happy
8	challenge	4	surprise	challenge	X	X	challenge	X
9	challenge	4	challenge	challenge	surprise	challenge	scared	sad
10	bored	2	challenge	annoyed	X	X	challenge	happy
11	challenge	5	happy	sad	challenge	X	surprise	happy
12	challenge	3	X	annoyed	stress	frustration	stress	happy
13	challenge	5	happy	annoyed	X	challenge	challenge	happy
14	effort	3	happy	annoyed	challenge	scared	X	happy
15	challenge	4	challenge	sad	scared	X	challenge	happy
16	effort	4	scared	frustration	X	frustration	scared	happy
17	effort	3	X	X	scared	X	fear	happy
18	challenge	5	challenge	X	scared	X	surprise	happy
19	stress	2	bored	angry	effort	X	angry	happy
20	fear	3	surprise	annoyed	fear	stress	challenge	happy
21	effort	3	fear	X	challenge	angry	fear	happy
22	challenge	4	challenge	annoyed	bored	X	fear	happy
23	stress	3	happy	X	challenge	X	challenge	happy
24	challenge	4	surprise	X	X	X	X	X
25	challenge	3	challenge	frustration	X	X	happy	sad

Table B.6.: Self-Assessment: **Cliff Scene**

ID	Treehouse	Fun	Mountain	Fun	Spider
1	effort	0	effort	2	X
2	happy	2	happy	3	scared
3	challenge	4	angry	2	disgust
4	confused	3	scared	3	scared
5	effort	2	effort	2	confused
6	annoyed	1	bored	1	X
7	effort	4	effort	3	X
8	X	1	annoyed	2	X
9	challenge	2	X	3	scared
10	happy	4	bored	2	scared
11	happy	5	effort	4	surprise
12	effort	2	effort	3	X
13	challenge	4	challenge	4	confused
14	happy	1	X	1	X
15	challenge	4	challenge	4	surprise
16	effort	2	bored	1	X
17	effort	3	effort	3	disgust
18	challenge	4	stress	3	scared
19	effort	1	scared	3	scared
20	challenge	4	challenge	4	fear
21	confused	1	effort	2	scared
22	bored	1	surprise	3	bored
23	effort	4	effort	3	scared
24	effort	5	confused	3	X
25	happy	5	challenge	5	surprise

Table B.7.: Self-Assessment: **Treehouse Scene** and **Mountain Scene**

ID	Body Feeling	Pleasure	Arousal	Dominance
1	6	9	6	5
2	8	9	6	5
3	7	8	4	6
4	6	8	4	6
5	6	9	5	5
6	7	7	7	6
7	7	7	4	5
8	6	7	3	5
9	8	8	6	7
10	7	8	7	7
11	7	8	5	8
12	3	7	3	4
13	9	8	7	7
14	8	5	7	6
15	4	7	3	3
16	5	6	5	6
17	7	6	4	4
18	7	7	7	5
19	8	7	5	7
20	8	9	5	6
21	9	7	4	6
22	3	8	4	8
23	7	6	5	6
24	8	9	7	7
25	9	7	6	8

Table B.8.: Self-Assessment prior to the Exergame

ID	Body Feeling	Pleasure	Arousal	Dominance	Borg
1	6	9	7	5	3
2	7	7	7	5	1
3	7	7	6	6	4
4	7	8	7	6	3
5	6	9	6	6	1
6	7	7	7	6	3
7	7	7	6	5	4
8	6	5	7	5	3
9	8	8	7	7	2
10	7	7	5	5	4
11	6	6	7	5	5
12	4	7	5	3	3
13	9	8	8	7	1
14	6	7	4	4	1
15	5	7	5	5	4
16	6	7	7	7	3
17	7	7	5	5	4
18	7	8	8	6	3
19	7	5	8	9	6
20	8	8	6	6	4
21	8	4	3	3	5
22	4	9	6	3	2
23	7	8	7	7	3
24	8	7	9	7	2
25	8	7	6	6	4

Table B.9.: Self-Assessment after the Journey

ID	Body Feeling	Pleasure	Arousal	Dominance	Borg
1	4	8	3	4	7
2	7	7	8	5	1
3	5	8	7	7	5
4	8	8	7	7	4
5	7	8	5	5	4
6	7	7	7	6	4
7	7	8	7	5	4
8	7	4	5	5	4
9	7	8	6	6	3
10	5	6	4	6	5
11	8	8	6	8	6
12	6	6	6	5	4
13	9	8	7	7	1
14	5	8	4	7	2
15	6	9	5	5	4
16	6	6	3	6	4
17	6	7	5	5	5
18	7	8	8	7	4
19	5	7	4	5	6
20	8	8	6	7	4
21	8	7	6	6	4
22	6	9	7	7	2
23	7	8	6	6	4
24	8	9	7	7	5
25	5	7	6	8	5

Table B.10.: Self-Assessment after the Exergame

C. Observer-Assessment

ID	Fork	Forest Scene	Saw Level	Coin Collection	Zombies
1	happy	X	X	effort	X
2	happy	scared	happy	effort	scared
3	happy	scared	happy	happy	scared
4	happy	scared	effort	X	scared
5	happy	scared	happy	happy	scared
6	surprise	scared	fear	happy	scared
7	X	X	X	X	scared
8	X	scared	fear	surprise	scared
9	happy	happy	X	happy	scared
10	happy	scared	happy	X	scared
11	effort	scared	happy	happy	scared
12	confused	scared	X	X	scared
13	confused	scared	X	happy	scared
14	X	happy	X	happy	happy
15	effort	scared	stress	X	scared
16	happy	scared	stress	happy	scared
17	happy	scared	happy	happy	scared
18	happy	scared	fear	scared	scared
19	confused	scared	fear	happy	scared
20	surprise	scared	X	X	scared
21	confused	surprise	happy	happy	surprise
22	happy	scared	X	happy	scared
23	happy	scared	effort	happy	scared
24	happy	scared	stress	effort	scared
25	effort	scared	stress	effort	scared

Table C.1.: Observer-Assessment: **Fork Scene** and **Forest Scene**

ID	Challenge	Falling	Boost	Landing	Teleport	Scene End
1	confused	confused	confused	happy	X	X
2	happy	X	X	happy	X	X
3	frustration	frustration	happy	happy	happy	happy
4	happy	X	confused	X	X	X
5	frustration	frustration	effort	happy	X	X
6	frustration	frustration	effort	happy	X	X
7	effort	X	X	X	X	X
8	frustration	frustration	confused	happy	X	happy
9	frustration	frustration	X	happy	X	happy
10	annoyed	frustration	X	X	happy	X
11	frustration	frustration	effort	happy	X	happy
12	X	frustration	surprise	X	frustration	frustration
13	challenge	happy	effort	happy	X	X
14	challenge	frustration	effort	happy	X	happy
15	frustration	frustration	annoyed	surprise	X	happy
16	frustration	frustration	happy	X	sad	sad
17	challenge	frustration	happy	X	X	X
18	challenge	happy	effort	happy	X	happy
19	frustration	frustration	effort	happy	X	happy
20	X	frustration	surprise	happy	X	happy
21	happy	X	effort	surprise	X	X
22	frustration	frustration	effort	happy	X	happy
23	frustration	annoyed	effort	happy	X	happy
24	frustration	frustration	challenge	X	happy	X
25	frustration	frustration	effort	happy	X	happy

Table C.2.: Observer-Assessment: **Challenge Scene**

ID	Teddy	Saw Teddy	Teddy Hit
1	X	happy	happy
2	happy	X	happy
3	effort	happy	happy
4	X	happy	X
5	happy	scared	sad
6	happy	happy	happy
7	happy	happy	happy
8	happy	confused	happy
9	annoyed	happy	happy
10	happy	X	happy
11	effort	effort	happy
12	X	stress	surprise
13	effort	X	surprise
14	happy	happy	happy
15	effort	happy	confused
16	sad	confused	sad
17	happy	happy	happy
18	challenge	sad	X
19	happy	happy	happy
20	annoyed	X	disgust, sad
21	effort	X	disgust
22	bored	X	effort
23	happy	X	confused
24	sad	happy	sad
25	happy	confused	happy

Table C.3.: Observer-Assessment: **Teddy Scene**

ID	Frozensea	Coin	Snowmen	Ice	Falling	Scene End
1	X		X	confused	frustration	X
2	happy	happy	happy	challenge	happy	happy
3	scared	X	scared	confused	X	happy
4	challenge	happy	scared	challenge	frustration	happy
5	challenge	happy	scared		effort	happy
6	challenge	happy	scared	effort	X	X
7	effort	X	X	challenge	X	X
8	scared	happy	scared	challenge	frustration	happy
9	annoyed	surprise	surprise	X	frustration	
10	scared	challenge	scared	effort	annoyed	X
11	scared	effort	scared	challenge	surprise	X
12	effort	X	scared	confused, effort	frustration	fear
13	challenge	happy	scared	confused	X	X
14	challenge	effort	X	challenge	X	happy
15	effort	challenge	scared	confused	X	happy
16	annoyed	effort	scared	effort	frustration	happy
17	challenge	challenge	scared	challenge	frustration	happy
18	challenge	challenge	scared	challenge	X	happy
19	challenge	happy	effort	challenge	annoyed	happy
20	challenge	X	scared	confused	frustration	X
21	happy	happy	surprise	challenge	annoyed	X
22	challenge	happy	scared	confused	surprise	X
23	challenge	happy	scared	X	X	X
24	effort	effort	scared	confused	X	X
25	challenge	challenge	scared	effort	frustration	happy

Table C.4.: Observer-Assessment: **Frozensea Scene**

ID	Cliff	Saw Scene	Falling	Saw Rock	Rock Hit	Bridge	Scene End
1	annoyed	X	frustration	X	X	X	X
2	challenge	challenge	X	X	X	X	X
3	challenge	X	frustration	X	confused	challenge	X
4	challenge	X	sad	effort	X	challenge	X
5	challenge	challenge	frustration	surprise	challenge	challenge	X
6	frustration	fear	frustration	surprise	frustration	annoyed	happy
7	effort	X	frustration	X	X	X	X
8	effort	challenge	frustration	challenge	X	challenge	X
9	effort	X	frustration	X	X	scared	happy
10	effort	X	annoyed	effort	X	effort	X
11	challenge	effort	X	effort	X	fear	happy
12	fear	fear	frustration	fear	frustration	fear	happy
13	effort	challenge	frustration	challenge	confused	challenge	happy
14	challenge	X	frustration	X	annoyed	X	X
15	challenge	X	annoyed	X	X	X	X
16	stress	stress	frustration	fear	X	stress	happy
17	challenge	challenge	X	challenge	X	fear	X
18	challenge	stress	X	stress	X	stress	X
19	stress	annoyed	annoyed	stress	X	stress	happy
20	challenge	challenge	frustration	surprise	confused	challenge	happy
21	effort	effort	X	X	annoyed	fear	X
22	challenge	challenge	frustration	challenge	X	fear	happy
23	effort	X	confused	X	X	challenge	happy
24	effort	happy	X	stress	X	X	X
25	effort	challenge	frustration	challenge	X	frustration	X

Table C.5.: Observer-Assessment: **Cliff Scene**

ID	Treehouse	Mountain	Spider
1	effort	effort	X
2	X	effort	scared
3	effort	effort	disgust
4	happy	scared	scared
5	effort	effort	confused
6	effort	effort	scared
7	challenge	effort	X
8	effort	X	X
9	happy	X	X
10	effort	effort	surprise
11	challenge	challenge	surprise
12	effort	X	surprise
13	effort	effort	X
14	happy	X	X
15	effort	challenge	X
16	effort	effort	X
17	effort	effort	fear
18	effort	scared	scared
19	effort	scared	scared
20	challenge	challenge	surprise
21	effort	surprise	surprise
22	bored	bored	X
23	effort	effort	surprise
24	effort	effort	X
25	effort	effort	surprise

Table C.6.: Observer-Assessment: **Treehouse Scene** and **Mountain Scene**