

Kristina Schaaff

# Integration von physiologischem Feedback in Lernanwendungen unter Alltagsbedingungen



# Integration von physiologischem Feedback in Lernanwendungen unter Alltagsbedingungen

Zur Erlangung des akademischen Grades eines

DOKTOR-INGENIEURS

von der Fakultät für Elektrotechnik und Informationstechnik  
des Karlsruher Instituts für Technologie (KIT)  
genehmigte

DISSERTATION

von

Dipl.-Inform. Kristina Schaaff  
geb. in Braunschweig

Hauptreferent: Prof. Dr. rer. nat. Wilhelm Stork  
Korreferent: Prof. Dr. Gerd Gidion  
Tag der mündlichen Prüfung: 21. Juli 2014



# Danksagung

Diese Dissertation entstand im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin am FZI Forschungszentrum Informatik. Für die Möglichkeit der Promotion möchte ich mich insbesondere bei meinem Doktorvater Prof. Dr. rer. nat. Wilhelm Stork bedanken, ohne dessen langjährige Unterstützung diese Arbeit nicht möglich gewesen wäre. Außerdem möchte ich Prof. Dr. Gerd Gidion für die Übernahme des Korreferats und die interessanten Diskussionen über die interdisziplinären Aspekte meiner Arbeit danken.

Ganz besonders möchte ich mich an dieser Stelle auch bei Dr. Marc Adam für seine Unterstützung bei der Durchführung zahlreicher Studien sowie die hiermit verbundenen Diskussionen zur Auswertung der Daten bedanken.

Meinen Kollegen und ehemaligen Kollegen danke ich für die angenehme und motivierende Arbeitsatmosphäre. Insbesondere Lars, Christoph und Malte möchte ich für die vielen konstruktiven Gespräche und Ratschläge zu meiner Arbeit danken. Stephan danke ich für das Rücken freihalten in der stressigen Endphase. Nora dafür, dass sie immer an mich geglaubt hat. Zudem möchte ich allen von mir betreuten Studenten danken, die durch ihre Arbeiten wichtige Beiträge zu vielen Aspekten des entwickelten Gesamtsystems geleistet haben. Ein großer Dank geht auch an Birgit Voss, die sich die Zeit genommen hat, meine Arbeit vollständig Korrektur zu lesen sowie an Pawel Bolek für die Gestaltung des Umschlags.

Zuletzt möchte ich mich natürlich auch bei meinen Eltern Ilona und Ehrenfried bedanken, die sowohl während des Studiums als auch während der anschließenden Promotion immer hinter mir standen und mir den nötigen Rückhalt gegeben haben.

Karlsruhe, 30. Juli 2014





# Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Eigener Beitrag	3
1.3	Aufbau der Arbeit	6
2	Grundlagen	7
2.1	Physiologische Signale und deren Analyse	7
2.1.1	Das menschliche Nervensystem	7
2.1.2	Das kardiovaskuläre System	8
2.1.3	Das elektrodermale System	14
2.1.4	Die Skelettmuskulatur	17
2.2	Emotionen und Lernen	20
2.2.1	Emotionen	21
2.2.2	Einfluss von Emotionen auf den Lernerfolg	23
2.2.3	Das Arbeitsgedächtnis	25
2.2.4	Emotionsregulation	29
2.3	Mathematische und algorithmische Grundlagen	30
2.3.1	Feature-Extraktion aus physiologischen Signalen	30
2.3.2	Maschinelles Lernen und Klassifikation	40
3	Stand der Technik	47
3.1	Systeme zur Aufzeichnung physiologischer Daten	47
3.2	Messung von Emotionen	49
3.2.1	Erfassung des emotionalen Erlebens	50
3.2.2	Analyse des beobachtbaren Verhaltens	51
3.2.3	Psychophysiologische Korrelate von Emotionen	52
3.3	Affective Computing	54
3.3.1	Anwendungsfelder	55
3.3.2	Klassifikation von Emotionen	55

3.3.3	Software-Lösungen zur Verarbeitung multimodaler Daten . . . . .	58
3.4	Technologiegestütztes Lernen . . . . .	58
3.4.1	Affektive Lernumgebungen . . . . .	59
3.4.2	Training des Arbeitsgedächtnisses . . . . .	61
3.4.3	Emotionsregulationstraining . . . . .	63
4	Systemkonzeption und -design . . . . .	65
4.1	Problemstellung und Ansatz . . . . .	65
4.1.1	Anforderungen an die Sensorschicht . . . . .	68
4.1.2	Anforderungen an die Datenverarbeitungsschicht . . . . .	69
4.1.3	Anforderungen an die Interaktionsschicht . . . . .	70
4.2	Auswahl geeigneter Biosignale und Messsysteme . . . . .	70
4.2.1	Kommerzielle Systeme zur Messung physiologischer Signale . . . . .	72
4.2.2	Entwicklung einer Maus zur Erfassung des Tastendrucks . . . . .	74
4.2.3	Einordnung der verwendeten Messsysteme . . . . .	78
4.3	xAffect – Ein Software-Framework für die Online-Analyse physiologischer Daten . . . . .	79
4.3.1	Verknüpfung von Komponenten . . . . .	81
4.3.2	Lebenszyklus der Komponenten . . . . .	82
4.3.3	Datenstrukturen . . . . .	84
4.3.4	Steuerung der xAffect-Software . . . . .	85
4.3.5	Integrierte Komponenten . . . . .	86
4.4	Lernanwendungen zur Interaktion mit dem Nutzer . . . . .	89
5	Erkennung von Erregung aus physiologischen Signalen . . . . .	91
5.1	Studiendesign . . . . .	91
5.1.1	Studienaufbau . . . . .	92
5.1.2	Studienablauf . . . . .	95
5.2	Datenanalyse . . . . .	96
5.2.1	Vorbereitung der Daten . . . . .	98
5.2.2	Vorabbetrachtungen . . . . .	102
5.2.3	Analyse der kardiovaskulären Aktivität . . . . .	107
5.2.4	Analyse der EDA . . . . .	111
5.2.5	Analyse der auf die Maus ausgeübten Kraft . . . . .	115

5.2.6	Multimodale Klassifikation . . . . .	117
5.3	Übertragbarkeit der Ergebnisse . . . . .	120
5.3.1	Studiendesign . . . . .	120
5.3.2	Datenanalyse . . . . .	121
5.4	Zusammenfassung der Ergebnisse . . . . .	126
6	Eignung von EDA für die kurzfristige Erregungserkennung	129
6.1	Studiendesign . . . . .	129
6.1.1	Studienaufbau . . . . .	129
6.1.2	Studienablauf . . . . .	130
6.2	Datenanalyse . . . . .	132
6.2.1	Vorbereitung der Daten . . . . .	132
6.2.2	Erkennung von EDA-Reaktionen . . . . .	133
6.2.3	Wirkung emotionaler Stimuli auf EDA und Herzrate	139
6.3	Zusammenfassung der Ergebnisse . . . . .	141
7	Nutzung von physiologischem Feedback zum Emotionsregulationstraining	143
7.1	Studiendesign . . . . .	143
7.1.1	Studienaufbau . . . . .	144
7.1.2	Studienablauf . . . . .	148
7.2	Datenanalyse . . . . .	150
7.2.1	Einfluss der Spielschwierigkeit auf die Performance	151
7.2.2	Einsatz von Emotionsregulation während des Spiels	152
7.2.3	Einfluss des ER-Trainings auf die Spielperformance	156
7.2.4	Langfristige Veränderungen des Emotionsregulationsverhaltens . . . . .	158
7.2.5	Transfer-Effekte auf andere Aufgaben . . . . .	159
7.3	Zusammenfassung der Ergebnisse . . . . .	163
8	Integration von physiologischem Feedback in ein Arbeitsgedächtnistraining	165
8.1	Anwendungsdesign . . . . .	165
8.1.1	Trainingsprinzip . . . . .	166
8.1.2	Integration des physiologischen Feedbacks . . . . .	167
8.1.3	Konzeptionsstudie zur Feedbackanzeige . . . . .	168
8.1.4	Entwicklung des Adaptionalgorithmus . . . . .	170

## Inhaltsverzeichnis

---

8.2	Systemanalyse . . . . .	178
8.2.1	Adaptionsverhalten der Anwendung . . . . .	179
8.2.2	Analyse des Trainingserlebnisses . . . . .	183
8.2.3	Analyse der Performanceveränderung . . . . .	184
8.2.4	Untersuchung der Systemusability . . . . .	184
8.3	Zusammenfassung der Ergebnisse . . . . .	185
9	Zusammenfassung und Ausblick	187
9.1	Zusammenfassung . . . . .	187
9.2	Ausblick . . . . .	189
	Literaturverzeichnis	191
	Eigene Veröffentlichungen	211
	Betreute studentische Arbeiten	213
	Abbildungsverzeichnis	215
	Tabellenverzeichnis	219
	Abkürzungsverzeichnis	221
A	Werte der HRV-Parameter	223

# 1 Einleitung

Der Einfluss von Emotionen auf den Lernerfolg ist schon lange bekannt und hat bereits festen Einzug in pädagogische Konzepte gefunden. Die Integration von Informationen über den emotionalen Zustand des Lernenden in technologiegestützte Lernumgebungen hingegen stellt eine noch recht junge Forschungsdisziplin dar. Erste Überlegungen zu diesem Thema reichen weniger als 20 Jahre zurück. Unter technologiegestütztem Lernen sind Lern- und Lehrsituationen zu verstehen, bei welchen Technologien wie Computer oder Sensoren eingesetzt werden (Dror, 2008). Durch die immer größer werdende Allgegenwärtigkeit technologischer Systeme in unserem Alltag kommt dem technologiegestützten Lernen eine stetig wachsende Bedeutung zu (Chan et al., 2006). Mit den technologischen Entwicklungen geht auch eine wachsende Mobilität der Lernsysteme einher. Hierdurch findet Lernen immer häufiger unter veränderlichen Umgebungsbedingungen statt, was oftmals Veränderungen des emotionalen Zustands einer Person bedingt. Die Berücksichtigung des emotionalen Zustands einer Person wird somit umso wichtiger, um einen optimalen Lernerfolg erzielen zu können.

## 1.1 Motivation

Derzeit sind in den Medien rasante Entwicklungen hinsichtlich mobiler Sensortechnologien zu beobachten. Nahezu wöchentlich gibt es neue Meldungen über neuartige Sensoren, die der Erfassung von Informationen über den Zustand und die Aktivitäten eines Nutzers dienen. Der aktuelle Trend zur Quantifizierung des Menschen – welcher sich in zahlreichen Quantified Self-Technologien widerspiegelt – ermöglicht die konstante Erfassung von Nutzerinformationen mittels speziell für diese Zwecke entwickelter Sensoren zu jeder Zeit und an jedem Ort. Dies wäre noch vor wenigen Jahren undenkbar gewesen. Hierdurch eröffnen sich neue Anwendungsfelder für psychophysiologisches Monitoring, da dieses dank der neuen Technologien

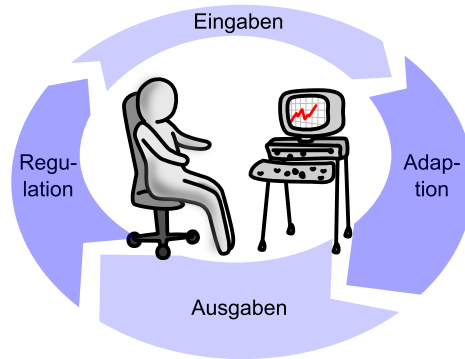
nun auch außerhalb von Laborumgebungen genutzt werden kann. Eines dieser Anwendungsfelder stellen affektive Lernumgebungen dar, bei welchen Informationen über den emotionalen Zustand einer Person mittels Sensorik erfasst und in den Lernprozess integriert werden.

Die Effizienz herkömmlicher computerbasierter Lernsysteme reicht bisher bei weitem nicht an die Lerneffizienz des 1:1-Unterrichts zwischen Lehrer und Lernenden heran. Dies kann auf die fehlende Berücksichtigung des emotionalen Nutzerzustands zurückgeführt werden (du Boulay et al., 2010), welche beim 1:1-Unterricht selbstverständlich ist. Affektive Lernsysteme versuchen diesem Faktor Rechnung zu tragen: Werden Informationen über den emotionalen Zustand einer Person in den Lernprozess integriert, kann dies dazu beitragen, den Lernenden emotional und motivational in diesem Prozess zu unterstützen (Antonacopoulou & Gabriel, 2001).

Gemäß des so genannten Yerkes-Dodson-Gesetzes (Yerkes & Dodson, 1908) gibt es für das Lernen einen optimalen emotionalen Erregungszustand, bei welchem eine maximale Leistung erzielt wird. Ähnliche Annahmen finden sich auch in Csikszentmihalyis Flow-Modell, nach welchem der Flow-Zustand erreicht wird, wenn Herausforderungen und Fähigkeiten möglichst gut übereinstimmen (Csikszentmihalyi & Csikszentmihalyi, 1988). Übersteigen die Anforderungen die Fähigkeiten eines Nutzers, führt dies zu einem zu hohen Erregungsniveau während Unterforderung ein zu niedriges Erregungsniveau bedingt. Umso näher sich der Erregungszustand eines Nutzers dem optimalen Zustand annähert, desto effizienter kann demnach das Lernen erfolgen.

Derzeitige affektive Lernsysteme zielen in erster Linie auf die Vermittlung deklarativen Wissens ab. Im Lernprozess wird hierbei versucht, die Kommunikation zwischen Mensch und System möglichst menschenähnlich zu gestalten. Informationen über den emotionalen Zustand des Nutzers werden aus physiologischen Signalen gewonnen.

Doch auch neben der Vermittlung deklarativen Wissens gibt es weitreichende Möglichkeiten physiologisches Feedback in den Lernprozess zu integrieren. Betrachtet man technologiegestütztes Lernen als Regelkreis, so ergibt sich das in Abbildung 1.1 dargestellte Bild. Seitens des Lernenden erfolgen Eingaben in das Lernsystem. Diese können zum einen inhaltlicher Natur sein, zum anderen jedoch auch Kontextinformationen über den Lernenden beinhalten – so zum Beispiel Informationen über dessen emotionalen Zustand. Seitens des Lernsystems erfolgt die Vermitt-



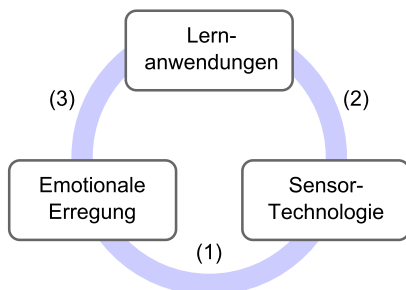
**Abbildung 1.1:** Regelkreis des technologiegestützten Lernens

lung von Lerninhalten an den Lernenden (Ausgaben). Für ein optimales Lernergebnis gibt es in diesem Regelkreis zwei Möglichkeiten, wie eine Anpassung des Gesamtsystems zu einem optimalen Erregungszustand des Lernenden führen kann. Das erste Stellglied ist hierbei die Lernanwendung. In dieser kann das Tempo der Stoffvermittlung oder auch der Anspruch des Lernstoffs angepasst werden (Adaption). Das zweite Stellglied stellt der Lernende selbst dar. Mittels entsprechender Techniken kann dieser versuchen, sein emotionales Erregungsniveau optimal an die Anforderungen der Lernanwendung anzupassen (Regulation).

## 1.2 Eigener Beitrag

Ziel dieser Arbeit ist die Erfassung von emotionaler Erregung mit Hilfe von psychophysiologischer Sensorik sowie die Integration dieser Informationen in Lernanwendungen unter Alltagsbedingungen. Hierbei erfolgt eine Verknüpfung unterschiedlicher Forschungsgebiete, welche in Abbildung 1.2 dargestellt sind. Der Fokus dieser Arbeit liegt auf der Erkennung emotionaler Erregung aus physiologischen Daten (1) sowie deren Integration in Lernanwendungen (2). Der Einfluss von emotionaler Erregung auf das Lernverhalten (3) wurde bereits hinreichend erforscht und ist daher nicht Bestandteil dieser Arbeit.





**Abbildung 1.2:** Verknüpfte Forschungsgebiete im Rahmen dieser Arbeit

Es wird ein System vorgestellt, welches mittels Sensortechnologien Informationen über den emotionalen Erregungszustand einer Person erfasst (1). Der Schwerpunkt des Systems liegt auf der möglichst belastungsfreien Messbarkeit dieser Signale. Neben kommerziell verfügbarer Sensorik wurde im Rahmen dieser Arbeit eine Maus entwickelt, welche in der Lage ist, die Krafteinwirkung auf die Maustaste zu erfassen und somit Rückschlüsse auf den Erregungszustand einer Person zu ermöglichen. Die Integration der Sensorik erfolgte hierbei derart, dass die Manipulation für den Nutzer nicht sichtbar ist.

Basierend auf physiologischen Sensordaten wurde ein Verfahren zur multimodalen Detektion emotionaler Erregung entwickelt. Die Entwicklung des Verfahrens erfolgte so, dass die Erregung mit einer möglichst geringen Latenz erkannt werden kann. Hierzu wurde erstmals auch die Eignung von Herzratenvariabilität für die Erregungserkennung mit kürzeren Zeitfenstern als den in Malik (1996) vorgegebenen Längen untersucht. Zudem wurde das Verfahren so entwickelt, dass es flexibel auf das Ausfallen einzelner Datenkanäle reagieren kann. Zusätzlich wurde basierend auf der elektrodermalen Aktivität (EDA) ein Verfahren zur Identifikation kurzfristiger emotionaler Reize unter Alltagsbedingungen entwickelt. Hierzu wurden auch Kontextinformationen aus körperlicher Aktivität und Herzaktivität hinzugenommen.

Für die Einbindung physiologischer Sensordaten in Lernanwendungen (2) wurde eine Middleware entwickelt, welche es ermöglicht, die physiologischen Sensordaten in die entsprechenden Anwendungen zu integrieren. Der Fokus

lag hierbei auf einer hohen Modularität der Middleware, damit diese ohne großen Aufwand an neue Sensorlösungen und Algorithmen angepasst werden kann.

Im Gegensatz zu bisherigen Arbeiten bezüglich affektiver Lerntechnologien zielen die Anwendungen in dieser Arbeit nicht auf die Vermittlung konkreter Lerninhalte ab. Vielmehr werden Systeme vorgestellt und evaluiert, welche der Vermittlung von Fähigkeiten dienen, die für erfolgreiches Lernen erforderlich sind. Dies umfasst sowohl das Training der Emotionsregulationsfähigkeiten sowie das Training des Arbeitsgedächtnisses.

Die beiden dargestellten Anwendungsbeispiele verfolgen zwei unterschiedliche Ansätze. Während das Emotionsregulationstraining darauf abzielt, dem Nutzer Werkzeuge an die Hand zu geben, seinen Erregungszustand optimal an ein Training anzupassen, wird anhand des Arbeitsgedächtnistrainings eine Möglichkeit vorgestellt, wie ein System an den Nutzerzustand angepasst werden kann.

Das hier verwendete Emotionsregulationstraining ist einem Finanzmarkt-szenario nachempfunden, da erfolgreiche Emotionsregulation für die dortigen Akteure von großer Bedeutung ist. Physiologisches Feedback kommt hier zum Einsatz, um erfolgreiche Emotionsregulation zu belohnen bzw. zu bestrafen. Hierdurch soll der Nutzer lernen, seine Emotionen gezielter zu steuern.

Die Schwierigkeitsanpassung von Anwendungen zum Training des Arbeitsgedächtnisses wurde in bisherigen Arbeiten entweder gar nicht oder ausschließlich basierend auf der vorherigen Trainingsperformance vorgenommen. Die Möglichkeit, Informationen über den emotionalen Erregungszustand einer Person mittels physiologischem Feedback in das Training zu integrieren, wurde bisher nicht betrachtet. Aus diesem Grund wird im Rahmen dieser Arbeit ein Arbeitsgedächtnistraining vorgestellt, welches anhand der Informationen aus physiologischer Sensorik den Schwierigkeitsgrad derart adaptiert, dass der Lernende stets auf einem optimalen Erregungsniveau trainieren kann, bei welchem weder Über- noch Unterforderung vorliegt. Der vorgestellte Ansatz soll dazu dienen, die Trainings-schwierigkeit bereits anzupassen, bevor die Motivation eines Lernenden aufgrund von Frustration durch schlechte Performance sinkt. Genauso erfolgt eine Erhöhung der Trainingsschwierigkeit bereits, bevor dem Lernenden langweilig wird.

### 1.3 Aufbau der Arbeit

Diese Arbeit gliedert sich in zwei Teile. Im ersten Teil werden in Kapitel 2 und 3 die Grundlagen dargestellt, welche zum weiteren Verständnis dieser Arbeit erforderlich sind. Des Weiteren erfolgt eine Beschreibung des aktuellen Stands der Technik. Im zweiten Teil der Arbeit folgen die eigenen Arbeiten. In Kapitel 4 wird hierfür zunächst das Systemkonzept vorgestellt, auf welchem alle weiteren Kapitel basieren. In den darauf folgenden Kapiteln 5 und 6 werden Ergebnisse von Studien zur Evaluation verschiedener physiologischer Signale für die Erkennung von emotionaler Erregung vorgestellt. Im nächsten Schritt werden in Kapitel 7 und 8 zwei unterschiedliche Ansätze vorgestellt, wie die Informationen über den emotionalen Erregungszustand einer Person in Lernumgebungen integriert werden können. Die Arbeit schließt in Kapitel 9 mit einer übergreifenden Zusammenfassung der einzelnen Studienergebnisse und einem Ausblick auf zukünftige Arbeiten.

## 2 Grundlagen

Dieses Kapitel dient der Beschreibung der Grundlagen, welche für das Verständnis dieser Arbeit erforderlich sind. In Abschnitt 2.1 sind zunächst die physiologischen Grundlagen der im Rahmen dieser Arbeit verwendeten Signale dargestellt. Abschnitt 2.2 widmet sich anschließend den psychologischen Konzepten von Emotionen und Lernen. Im darauffolgenden Abschnitt wird dargestellt, wie sich Emotionen in physiologischen Signalen widerspiegeln. Abschließend werden die mathematischen Grundlagen dieser Arbeit erläutert.

### 2.1 Physiologische Signale und deren Analyse

Der nachfolgende Abschnitt erläutert, die physiologischen Grundlagen, welche im Rahmen dieser Arbeit von Bedeutung sind. Neben der grundlegenden Funktionsweise des menschlichen Nervensystems wird hierbei auf die kardiovaskuläre, die elektrodermale und die Muskelaktivität sowie die Messung der damit einhergehenden physiologischen Signale eingegangen.

#### 2.1.1 Das menschliche Nervensystem

Das menschliche Nervensystem besteht aus zwei Teilen: dem zentralen Nervensystem (ZNS) und dem peripheren Nervensystem (PNS), die eng miteinander verflochten sind. Das ZNS empfängt sensorische Informationen aus der Umgebung und gibt diese an das limbische System und den Kortex weiter. Das limbische System, das den Hypothalamus, den Hippocampus und die Amygdala beinhaltet, gilt als der Ursprung von Emotion und Motivation und ist ebenfalls verantwortlich für emotionale Reaktionen (Silbernagl & Despopoulos, 2007).

Das PNS wird in zwei wesentliche Bestandteile untergliedert: das somatische und das autonome Nervensystem (ANS). Das somatische Nervensys-

tem innerviert die Skelettmuskeln sowie die externen sensorischen Organe wie beispielsweise die Haut. Es wird oft auch als bewusstes Nervensystem bezeichnet, da die Reaktionen dieses Nervensystems mit Ausnahme von Reflexen der Skelettmuskulatur bewusst gesteuert werden können. Das ANS hingegen lässt sich nicht willentlich steuern. Es dient der Übermittlung von Impulsen aus dem ZNS zu den peripheren Organen. Ziel ist es, die Organe immer optimal an veränderte äußerliche Bedingungen anzupassen. Das ANS beinhaltet das sympathische und das parasympathische Nervensystem, welche den Erregungszustand des Körpers regulieren. In Abbildung 2.1 ist das Zusammenwirken des sympathischen und des parasympathischen Nervensystems dargestellt.

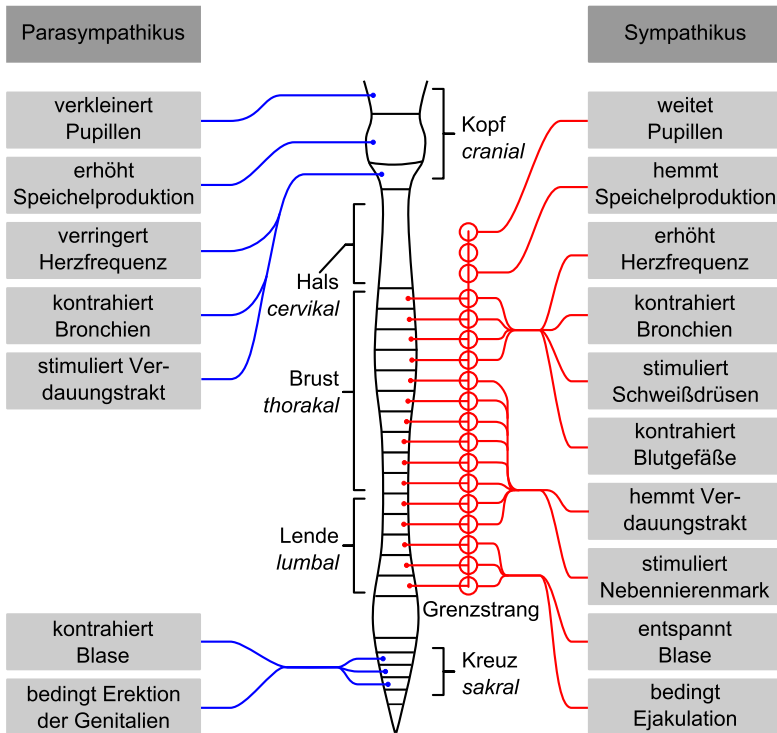
Wie man erkennen kann, werden durch das ANS eine Vielzahl physiologischer Signale (wie beispielsweise die Pulsfrequenz oder elektrodermale Aktivität) beeinflusst. Während das sympathische Nervensystem aktivierend (*'fight and flight'*) auf den Körper wirkt, übt das parasympathische Nervensystem einen regulierenden Einfluss (*'rest and digest'*) aus (Carlson, 2007).

### 2.1.2 Das kardiovaskuläre System

Im folgenden Abschnitt wird zunächst die Funktionsweise des kardiovaskulären Systems dargestellt. Im Anschluss daran folgt eine Erläuterung der Messung der elektrischen Herzaktivität, des Pulses sowie des Blutdrucks.

#### 2.1.2.1 Aufbau des Herz-Kreislauf-Systems

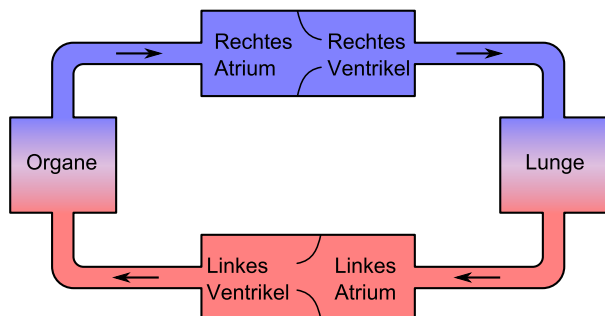
Das Herz-Kreislauf-System ist für die Stoffwechselforgänge des Körpers zuständig. Die zentrale Pumpe des Herz-Kreislauf-Systems stellt das Herz dar. Es ist dafür verantwortlich, Blut durch die Gefäße zu pumpen. Das Herz ist in zwei Hälften untergliedert, die sich aus je einem Vorhof (Atrium) und einer Herzkammer (Ventrikel) zusammensetzen. Die rechte Herzhälfte ist für die Versorgung der Lunge zuständig, während die linke Herzhälfte die Körperorgane sowie den Kopf mit Blut versorgt (Clauss, 2009). Der Aufbau des Herz-Kreislaufsystems ist in Abbildung 2.2 schematisch dargestellt. Im Körperkreislauf führen die vom Herz wegführenden Arterien den Organen sauerstoffreiches Blut zu. Die zum Herz hinführenden Venen transportieren das sauerstoffarme Blut zurück zum Herz. Im Lungenkreislauf hingegen



**Abbildung 2.1:** Zusammenwirken von sympathischem und parasympathischem Nervensystem

transportieren die Arterien das sauerstoffarme Blut zur Lunge, von wo aus es mit Sauerstoff angereichert wieder zurück zum Herz transportiert wird.

Der Herzzyklus besteht aus zwei Phasen: Während der *Systole* wird das Blut in die Arterien ausgeworfen, in der *Diastole* strömt das Blut aus den Venen zurück in das Herz. Die Diastole beginnt mit der Entspannung des Herzmuskels. Anschließend strömt das Blut aus den Vorhöfen in die Herzkammern ein. In der Systole baut daraufhin der Herzmuskel Spannung auf und das verbleibende Blut wird aktiv aus den Vorhöfen in die Kammer gepresst. In der darauf folgenden Auswurfphase kontrahiert der Herzmuskel



**Abbildung 2.2:** Schematische Darstellung des Herz-Kreislaufsystems. Rot kennzeichnet sauerstoffreiches Blut, blau sauerstoffarmes Blut

und das Blut wird aus dem Herz herausgepumpt. (Schandry, 1989)

Durch die stoßweisen Herzaktionen wird eine Pulswelle erzeugt, die sich entlang der arteriellen Strombahn ausbreitet. Die Strömungsgeschwindigkeit der Pulswelle ist abhängig vom Querschnitt der Blutgefäße in den jeweiligen Bereichen und nimmt mit sinkendem Querschnitt der Blutgefäße zu. Das strömende Blut übt einen ständigen Druck auf die Gefäße aus. Dieser Blutdruck variiert in Abhängigkeit von der Strömungsgeschwindigkeit des Blutes sowie dem Widerstand der Blutgefäße (Silbernagl & Despopoulos, 2007). Es wird unterschieden zwischen dem systolischen Blutdruck – dem Maximum des Blutdrucks in der Aorta während der systolischen Auswurfphase – und dem diastolischen Blutdruck, der dem Minimum des Drucks in der Aorta während der Diastole entspricht.

### 2.1.2.2 Messung der elektrischen Herzaktivität

Mit Hilfe der Elektrokardiographie wird die elektrische Aktivität des Herzens im Zeitverlauf gemessen. Das EKG-Signal spiegelt Potenzialdifferenzen auf der Körperoberfläche wider, die ihren Ursprung in der Herzerregung haben (Silbernagl & Despopoulos, 2007).

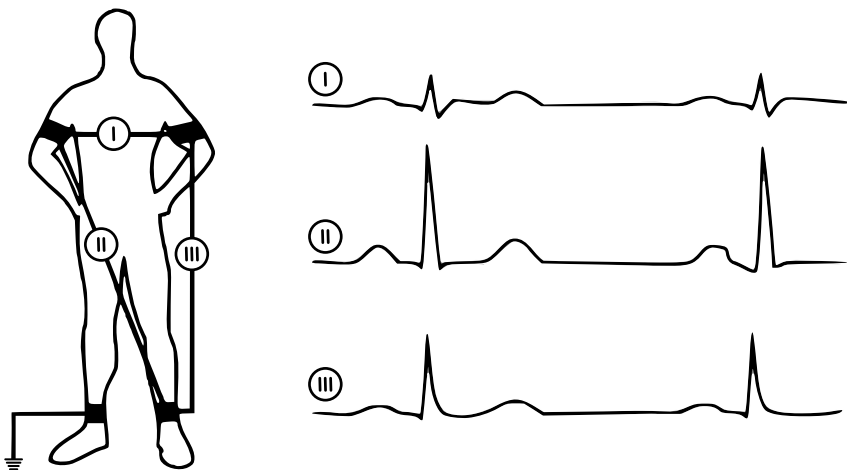
Für die Ableitung des EKG-Signals wird mindestens ein Elektrodenpaar auf dem Körper platziert. Insgesamt gibt es 12 verschiedene Standardableitungen. Die am häufigsten verwendete EKG-Standardableitungen nach Einthoven, bei welchen die Elektroden an den Extremitäten platziert

werden, sind in Abbildung 2.3 dargestellt. Wie aus der Abbildung ersichtlich wird, variiert die Morphologie des EKG-Signals in Abhängigkeit von der Elektrodenplatzierung.

Aufgrund der unterschiedlichen Signalverläufe in Abhängigkeit von der Elektrodenplatzierung ist es bei der Durchführung von Studien wichtig, dass die Elektroden immer an der selben Position platziert werden, wenn die Daten verschiedener Versuchspersonen oder aus verschiedenen Sitzungen verglichen werden sollen.

Das EKG-Signal besteht im Normalfall aus fünf aufeinanderfolgenden Elementen, die mit aufeinanderfolgenden Buchstaben gekennzeichnet sind (Silbernagl & Despopoulos, 2007):

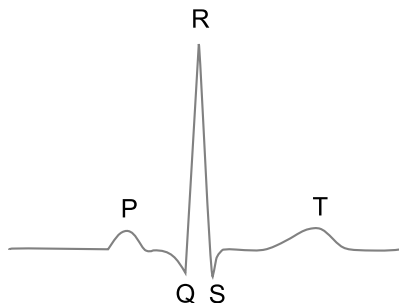
- Die **P-Welle** spiegelt die Depolarisation der Vorhöfe wieder und dauert üblicherweise zwischen 80 und 100 ms.
- Der **QRS-Komplex** – bestehend aus Q-Zacke, R-Zacke und S-Zacke – entspricht der Depolarisation der Kammer. Üblicherweise dauert der QRS-Komplex 60 bis 100 ms.
- Die **T-Welle** kennzeichnet die Repolarisation der Kammer.



**Abbildung 2.3:** EKG-Signal in Abhängigkeit von der Ableitungsposition (Gramann & Schandry, 2009)



Der Verlauf des EKG-Signals ist in Abbildung 2.4 dargestellt.



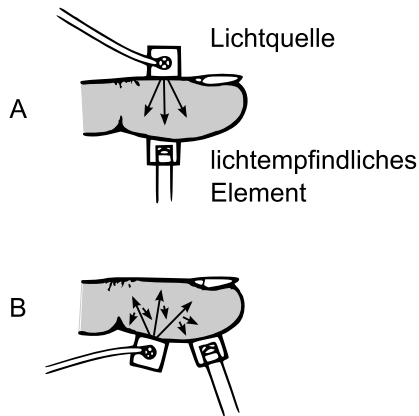
**Abbildung 2.4:** Typischer Verlauf eines EKG-Signals

Aufgrund ihrer spitzen Form ist die R-Zacke im EKG-Signal sehr leicht zu identifizieren. Aus diesem Grund wird diese üblicherweise zur Markierung einzelner Herzschläge herangezogen.

### 2.1.2.3 Pulsmessung

Bei der Photoplethysmographie (PPG) handelt es sich um ein optisches Verfahren zur Messung des Pulses. Das Prinzip der PPG beruht darauf, dass durchblutetes und undurchblutetes Gewebe unterschiedlich durchlässig für rotes Licht sind. Bei starker Durchblutung des Gewebes wird ein größerer Teil des Lichtes absorbiert, als bei geringer Durchblutung (Gramann & Schandry, 2009).

PPG-Sensoren bestehen aus einer lichtemittierenden Diode sowie einem Photosensor zur Messung der Lichtintensität. Es werden zwei verschiedene Arten von Messungen unterschieden: Bei der transmissiven PPG-Messung wird der Lichtanteil gemessen, der ein Gewebe durchstrahlt, während bei der reflexiven Messung der Lichtanteil gemessen wird, der von dem Gewebe reflektiert wird (siehe Abbildung 2.5). Bei den Transmissionsystemen liegen Lichtquelle und Photodiode einander gegenüber, während sich diese bei Reflexionssystemen beide auf einer Seite befinden. Transmissive Messungen werden üblicherweise an der Fingerspitze oder dem Ohrfläppchen durchgeführt. Bei der reflexiven Messung hingegen können Sensoren an fast allen Körperstellen platziert werden. In einer von Tobola & Douniama



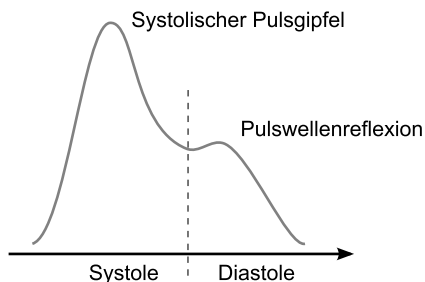
**Abbildung 2.5:** Arten von PPG-Messungen: A transmissiv, B reflexiv (Schandry, 1989)

im Jahr 2007 durchgeführten Studie wurde jedoch gezeigt, dass die beste Signalqualität an der Fingerspitze erzielt werden kann.

Abbildung 2.6 zeigt den typischen Verlauf einer Pulswelle. Während der Systole nimmt die Blutmenge in dem Gewebe stark zu, wodurch auch die Lichtabsorption steigt und die reflektierte bzw. durchstrahlende Lichtintensität abnimmt. Der Hochpunkt der Kurve kennzeichnet den systolischen Pulsgipfel. Im weiteren Signalverlauf ist eine zweite, kleinere Erhöhung der Pulswelle zu beobachten. Diese entsteht durch die Reflexion der Pulswelle beim Übergang in kleinere Blutgefäße.

### 2.1.2.4 Messung des Blutdrucks

Die Messung des Blutdrucks erfolgt in den meisten Fällen über das sogenannte Riva-Rocci-Verfahren mittels einer Druckmanschette, die um den Arm gelegt wird (Gramann & Schandry, 2009). Jedoch erlaubt dieses Verfahren keine kontinuierliche Messung des Blutdrucks. Zur kontinuierlichen, nichtinvasiven Messung des Blutdrucks kann die Puls-Transit-Zeit (engl. *pulse transit time*, PTT) herangezogen werden. Die PTT gibt die Zeit an, die die Pulswelle benötigt, um eine bestimmte Strecke im Körper zurückzulegen. Zur Bestimmung der PTT wird üblicherweise die Laufzeitdifferenz



**Abbildung 2.6:** Schematische Darstellung einer Pulswelle sowie typischer Kennwerte. Der systolische Pulsgipfel ist mit A gekennzeichnet, der zweite Gipfel entsteht durch die Reflexion der Pulswelle beim Übergang in die kleineren Blutgefäße.

zwischen einem EKG-Signal und dem an einer peripheren Körperstelle gemessenen PPG-Signal berechnet. Es ist zu beachten, dass die PTT in diesem Fall auch die Vorauswurfszeit (engl. *pre-ejection period*, PEP) beinhaltet. Die PEP ergibt sich aus der Verzögerung zwischen dem Einsetzen der elektrischen Herzaktivität und dem tatsächlichen Blutausswurf (Newlin, 1981). Die Zeit vom Beginn der PEP bis zur Ankunft der Pulswelle an einer peripheren Körperstelle wird daher oftmals auch als *pulse arrival time* (PAT) bezeichnet und setzt sich zusammen aus der PEP und der tatsächlichen PTT. Studien haben gezeigt, dass eine negative Korrelation zwischen systolischem Blutdruck und der PAT besteht wohingegen die Korrelation zum diastolischen und mittleren Blutdruck eher gering ist. Jedoch konnte eine hohe negative Korrelation von diastolischem und mittleren Blutdruck zu der tatsächlichen PTT nachgewiesen werden (z. B. Geddes et al., 1981; Lane et al., 1983; Payne et al., 2006).

Die Begriffe PAT und PTT werden oftmals synonym verwendet. Nachfolgend wird mit PTT die Zeit zwischen R-Zacke (und somit dem Beginn der PEP) und der Pulswelle bezeichnet.

### 2.1.3 Das elektrodermale System

Das elektrodermale System umfasst alle Veränderungen der elektrischen Eigenschaften der Haut. Diese Veränderungen werden auch als elektroder-

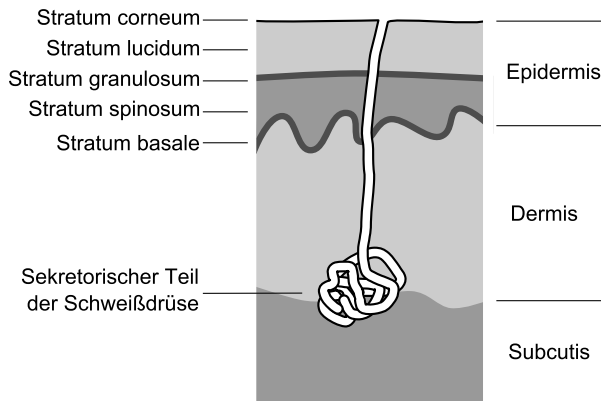
male Aktivität (EDA) bezeichnet und werden durch die Aktivierung von Schweißdrüsen bedingt. Nachfolgend wird zunächst der Aufbau der Haut und der Schweißdrüsen dargestellt, bevor anschließend auf die Messung der EDA eingegangen wird. Ein wesentlicher Unterschied des elektrodermalen Systems zum kardiovaskulären System ist, dass das elektrodermale System ausschließlich sympathisch innerviert ist (Dawson et al., 2007).

### 2.1.3.1 Aufbau der Haut und der Schweißdrüsen

Die Haut dient dem Schutz des Körperinneren vor äußeren Einflüssen wie beispielsweise dem Eindringen von Fremdkörpern oder Krankheitserregern. Auf der anderen Seite ermöglicht die Haut jedoch auch einen gezielten Stoffaustausch bei der Hautatmung und übernimmt eine wichtige Funktion bei thermoregulatorischen Prozessen durch Variation der Schweißproduktion (Dawson et al., 2007). Die Haut lässt sich in drei Untereinheiten gliedern: die Oberhaut (Epidermis) übernimmt die eigentliche Schutzfunktion der Haut. Sie besteht aus verhornten Epithelzellen und beinhaltet keine Blutgefäße. Darunter findet sich die aus Bindegewebsfasern bestehende Lederhaut (Dermis). Diese dient der Ernährung und Verankerung des Epithelgewebes der Epidermis und beinhaltet die Talgdrüsen. Im Gegensatz zur Epidermis beinhaltet sie auch feine Blutgefäße und glatte Muskulatur. Die Unterhaut (Subdermis) ist aus lockerem Bindegewebe und Fettgewebe aufgebaut, durch welches größere Blutgefäße verlaufen. In der Grenzschicht zwischen Leder- und Unterhaut finden sich die Schweißdrüsen (Gramann & Schandry, 2009). Es wird zwischen zwei Arten von Schweißdrüsen unterschieden: apokrine und ekkrine Drüsen. Die größeren apokrinen Drüsen findet man im Bereich der Achseln, den Brustwarzen und dem Genitalbereich, wohingegen die ekkrinen Drüsen über den gesamten Körper verteilt sind. Da die apokrinen Drüsen nicht zur elektrodermalen Aktivität beitragen, werden diese im weiteren Verlauf der Arbeit nicht betrachtet. Der Aufbau der ekkrinen Schweißdrüsen sowie deren Einbettung in die unterschiedlichen Hautschichten ist in Abbildung 2.7 dargestellt.

Bei den ekkrinen Schweißdrüsen kann zwischen verschiedenen Arten des Schwitzens unterschieden werden (Boucsein, 2011):

- *Thermoregulatorisches Schwitzen* ist für die Wärmeregulation des



**Abbildung 2.7:** Anatomie einer ekkrinen Schweißdrüse in den unterschiedlichen Hautschichten

Körpers zuständig.

- *Emotionales Schwitzen* wird durch Reaktionen auf psychologische oder emotionale Ereignisse ausgelöst.
- *Gustatorisches Schwitzen* ist bedingt durch sehr saures, salziges oder heißes Essen.
- *Allgegenwärtiges spontanes Schwitzen* kann im Zusammenhang mit dem entspannten Muskeltonus auftreten, jedoch ist dies noch nicht ausreichend belegt.
- *Reflexschwitzen* wird durch eine Veränderung des Umfelds (z.B. durch Strahlung oder Nadelstiche) an lokalen Stellen des Körpers hervorgerufen.
- *Pharmakologisches Schwitzen* kann beispielsweise durch Injektionen als auch durch Ionthophorese mit cholinergen Substanzen hervorgerufen werden.

Im Rahmen dieser Arbeit wird insbesondere das emotionale Schwitzen betrachtet. Im Gegensatz zu anderen Arten des Schwitzens tritt dieses in erster Linie an Händen und den Fußsohlen auf.

### 2.1.3.2 Messung der elektrodermalen Aktivität

Bei der Messung elektrodermalen Aktivität wird zwischen zwei verschiedenen Verfahren unterschieden:

- Bei dem *Konstantstromverfahren* wird ein konstanter Strom auf die Haut gegeben. Das erhaltene Signal gibt den Hautwiderstand an, der umso geringer wird, je mehr Schweißdrüsen aktiv sind.
- Das *Konstantspannungsverfahren* erfasst dagegen die elektrische Leitfähigkeit der Haut. Diese ändert sich ebenfalls in Abhängigkeit von der Schweißdrüsenaktivität: Je mehr Schweißdrüsen aktiv sind, desto höher ist auch die Hautleitfähigkeit.

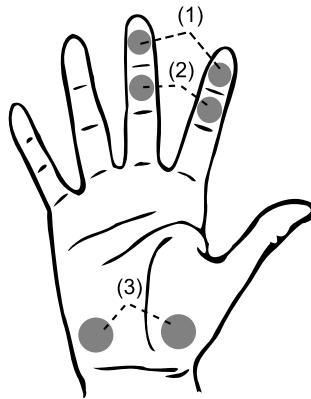
Die elektrische Leitfähigkeit wird in der Einheit Siemens ( $S$ ) angegeben, während der Hautwiderstand in Ohm ( $\Omega$ ) gemessen wird.

Bei der Messung von EDA ist es wichtig, nichtpolarisierende Elektroden zu verwenden, da die Polarisierung bei dem aufgenommenen Signal zu einem Drift führen würde. Aus diesem Grund werden in der Regel Ag/AgCl-Elektroden verwendet. Bei der Auswahl der Größe der Elektroden ist zu beachten, dass diese einerseits zwar so groß wie möglich sein sollten, auf der anderen Seite jedoch die von den Elektroden bedeckte Fläche eben sein sollte, was wiederum die mögliche Größe einschränkt. Um die Elektroden-Haut-Impedanz zu reduzieren, muss spezielles Elektrodengel verwendet werden, das im Gegensatz zu Elektrodengel wie es beispielsweise für EKG-Aufnahmen verwendet wird, eine ähnliche Elektrolytkonzentration wie die Haut aufweist. Für eine genauere Beschreibung des Elektrodengels sei auf Fowles et al. (1981) verwiesen.

In der Regel wird für Messungen der EDA die Handfläche verwendet, da diese eine hohe Dichte an aktiven Schweißdrüsen aufweist. Abbildung 2.8 veranschaulicht die Standard-Ableitpositionen, die an der Hand verwendet werden. Ist eine Messung an der Hand nicht möglich, können alternativ auch die Fußsohlen verwendet werden. Weitere mögliche Ableitpositionen werden von Rickles & Day (1968) diskutiert.

### 2.1.4 Die Skelettmuskulatur

Die Skelettmuskulatur ist für die willkürlichen Bewegungen des menschlichen Körpers verantwortlich. Nachfolgend wird daher zunächst der Aufbau



**Abbildung 2.8:** Standard-Positionen für EDA-Ableitungen an der Handfläche mit Elektrodenplatzierungen: (1) an den oberen Fingergliedern, (2) den mittleren Fingergliedern oder (3) am Daumen- und Kleinfingerballen

der Skelettmuskulatur erläutert und anschließend auf die Messung der Kraft, die durch die Muskulatur ausgeübt wird eingegangen.

### 2.1.4.1 Aufbau der Skelettmuskulatur

Die Skelettmuskulatur setzt sich aus einer großen Anzahl quergestreifter Muskeln zusammen. Ein quergestreifter Muskel wiederum besteht aus mehreren Muskelfasern. Die Muskelfasern enden an den Sehnen, welche die Kontraktionskräfte der Muskeln auf das Skelett übertragen. Jede Muskelfaser wird von einem Motoneuron innerviert, wobei ein Motoneuron mehrere Muskelfasern innervieren kann. Ein Motoneuron mit den dazugehörigen Muskelfasern wird auch als *motorische Einheit* bezeichnet. Die Zahl der Muskelfasern, die von einem Motoneuron innerviert wird, schwankt je nach Aufgabe des jeweiligen Muskels. Während bei den Augenmuskeln weniger als 10 Fasern von einem Motoneuron versorgt werden, kann die Anzahl bei Muskeln zur Ausführung grober Körperbewegungen bis zu 1000 Fasern betragen. (Schandry, 1989)

Die Stärke der Muskelaktivität kann durch verschiedene Faktoren be-

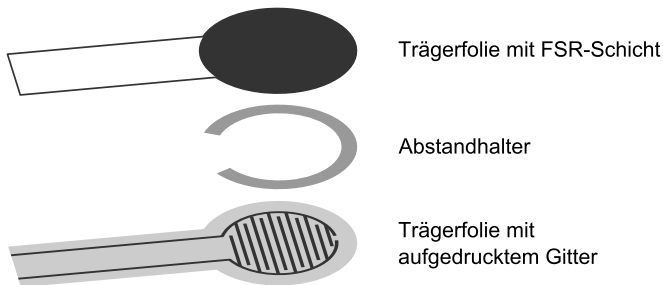
einflusst werden: durch die Anzahl der aktivierten motorischen Einheiten sowie durch die Erhöhung der Entladungsfrequenz der Neuronen. Je mehr motorische Einheiten ein Muskel enthält, desto feiner kann auch die entsprechende Bewegung abgestuft werden (Silbernagl & Despopoulos, 2007).

### 2.1.4.2 Messung der Muskelkraft

Die von einem Muskel ausgeübte Kraft wird im Rahmen dieser Arbeit mit Hilfe von Kraftsensoren gemessen. Im Rahmen dieser Arbeit wurden zwei verschiedene Arten von Kraftsensoren verwendet: Force Sensing Resistors (FSR) sowie piezoresistive Sensoren.

#### Force Sensing Resistors

Bei FSRs handelt es sich um einen mehrschichtigen Sensor, dessen Leitfähigkeit mit zunehmendem Druck ansteigt. Die untere Schicht besteht aus zwei kammförmig ineinander verschachtelten, elektrisch leitenden Kontaktgittern. Die obere Schicht verfügt über eine halbleitende Beschichtung. Beide Schichten sind durch einen Abstandhalter voneinander getrennt, so dass sie sich nicht berühren. Abbildung 2.9 veranschaulicht den Aufbau eines FSRs.



**Abbildung 2.9:** Aufbau eines FSRs

Wird eine Kraft auf den Drucksensor ausgeübt, kommt es zu einem Kontakt der halbleitenden Schicht mit den Kontaktgittern. Je größer die Kontaktfläche zwischen den beiden Schichten, umso mehr nimmt auch die elektrische Leitfähigkeit des Sensors zu (Fraden, 2004).



### Piezoresistive Kraftmessung

Bei Krafteinwirkung auf einen metallischen Leiter kommt es zu einer Verformung des Materials. Diese Verformung hat eine Veränderung des spezifischen elektrischen Widerstands zur Folge.

Bei einem metallischen Leiter der Länge  $l$  mit einem Querschnitt  $A$  und einem spezifischen elektrischen Widerstand  $\rho$  kann der elektrische Widerstand  $R$  wie folgt berechnet werden:

$$R = \rho \frac{l}{A} \quad (2.1)$$

Der Widerstand wird somit mit zunehmender Länge und abnehmendem Querschnitt – wie es bei einer Dehnung der Fall ist – größer.

Die relative Widerstandsänderung in Abhängigkeit von der Längsdehnung leitet sich aus folgender Gleichung ab:

$$\frac{\Delta R}{R} = k \cdot \frac{\Delta l}{l} \quad (2.2)$$

wobei der  $k$ -Faktor ein Maß für die Empfindlichkeit des Dehnmessstreifens (DMS) ist. Bei Halbleiter-DMS ist der  $k$ -Faktor aufgrund der Dotierung und Ausrichtung der Kristalle um ein Vielfaches größer als bei Metall-DMS. Dies führt dazu, dass Widerstandsänderungen auch ohne nennenswerte Veränderungen der Geometrie gemessen werden können. Die Veränderung des elektrischen Widerstands des Materials durch Druck oder Zug, welche nicht auf die Dehnung zurückgeführt werden kann, wird als *piezoresistiver Effekt* bezeichnet (Hering & Schönfelder, 2011).

Der Vorteil bei der piezoresistiven Kraftmessung gegenüber Metall-DMS ist insbesondere in der weitaus höheren Messgenauigkeit zu sehen.

## 2.2 Emotionen und Lernen

Emotionen sind ein wesentlicher Bestandteil des menschlichen Verhaltens und beeinflussen uns in unserem Handeln. Auch auf das Lernen üben sie einen entscheidenden Einfluss aus. Dieser Abschnitt soll die Grundlagen legen, die für das Verständnis von Emotionen und deren Einfluss auf den Lernerfolg notwendig sind. Zusätzlich wird das Konzept des Arbeits-

gedächtnisses eingeführt, welches maßgeblich an Lernprozessen beteiligt ist.

### 2.2.1 Emotionen

Bis heute gibt es keine einheitliche Definition des Begriffs der Emotion. Eine sehr treffende Aussage zu dieser Thematik findet sich bei Russell & Fehr (1984):

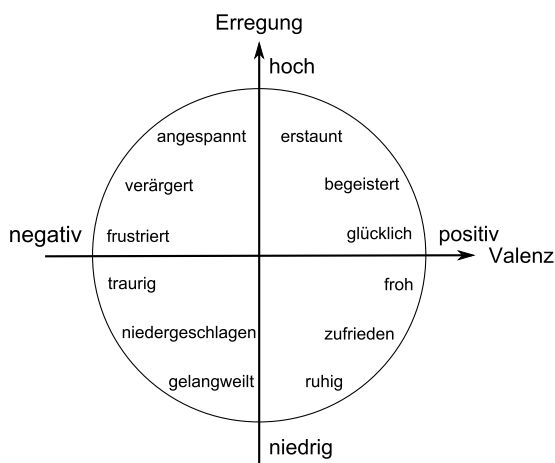
“ Everyone knows what an emotion is, until asked to give a definition. ”

Als die beiden grundlegenden Richtungen zur Einordnung von Emotionen sind kategoriale Modelle und dimensionale Modelle zu nennen. Während bei kategorialen Modellen von einem klar definierten Set an sogenannten Basisemotionen ausgegangen wird (siehe hierzu z. B. Ekman, 1992), werden Emotionen in den dimensional Modellen als Punkte in einem mehrdimensionalen Raum dargestellt. Der Vorteil dieser Darstellungsweise liegt darin, dass die Emotionen im Gegensatz zu den kategorialen Modellen nicht innerhalb vorgegebener Grenzen eingeordnet werden müssen, sondern vielmehr ein kontinuierlicher Raum innerhalb der Emotions-Dimensionen zur Einordnung zur Verfügung steht.

Die Ursprünge der dimensional Darstellungsweise gehen zurück auf Spencer (1890), der Emotionen als Dimensionen der Bewusstseinsphäre beschrieb. Diese Theorie wurde 1896 von Wundt durch Benennung dreier bipolarer Dimensionen erweitert: (1) angenehm – unangenehm, (2) Entspannung – Anspannung sowie (3) Ruhe – Aufregung. Eine empirische Validierung dieser Dimensionen mittels der Faktoranalyse erfolgte 1952 durch Osgood. Basierend auf diesen Ergebnissen wurde von Bradley & Lang (1994) ein Modell basierend auf den drei Dimensionen Erregung, Valenz und Dominanz eingeführt, das bis heute Bestand hat. Die Dimensionen sind wie folgt definiert:

- *Erregung*: quantitativer Grad der Aktivierung (ruhig ↔ aufgeregt)
- *Valenz*: Qualität der Emotion (negativ ↔ positiv)
- *Dominanz*: Grad der Kontrolle, den eine Person über eine Situation hat (schwach ↔ stark)

Die am häufigsten verwendeten Dimensionen stellen hierbei Erregung und Valenz dar. Ergebnisse von Russel & Mehrabian (1977); Russel (1979, 1980) belegen, dass diese Dimensionen den Großteil der Varianz auf emotionalen Skalen aufklären, während Dominanz einen vergleichsweise kleinen Einfluss hat. Auch das von Russel (1980) geprägte *Circumplex model of affect* stellt Emotionen anhand ihrer Position in dem zweidimensionalen Raum dar, welcher von den Achsen Valenz und Erregung aufgespannt wird. Abbildung 2.10 stellt beispielhaft die Einordnung von Emotionen in den zweidimensionalen Raum dar.



**Abbildung 2.10:** Exemplarische Darstellung des zweidimensionalen Emotionsmodells

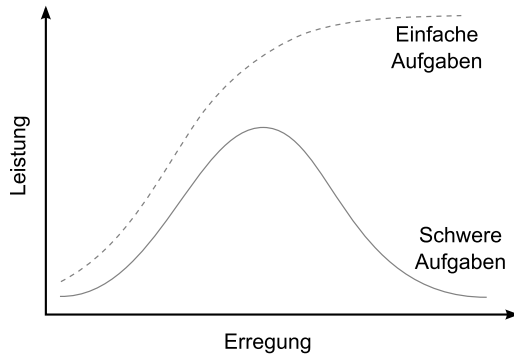
Die bedeutendste Emotion für die Lernperformance stellt die emotionale Erregung dar (siehe hierzu Abschnitt 2.2.2). Gemäß des dimensionalen Emotionsmodells gibt Erregung den Grad der emotionalen Aktiviertheit an. Schwache emotionale Erregung entspricht dabei einem ruhigen, entspannten Zustand, während starke Erregung mit einer hohen Aktiviertheit verbunden ist. Von Stress grenzt sich emotionale Erregung dadurch ab, dass sich Stress insbesondere im Arbeitsumfeld als „ein negatives emotionales Erlebnis, das mit Empfindungen von Angst, Spannung, Niedergeschlagenheit, Ärger, Müdigkeit und von mangelnder Tatkraft verbunden ist“ (Grandjean, 1991)

darstellt. Im Gegensatz zu emotionaler Erregung werden bei Stress somit nur Zustände berücksichtigt, die mit einer negativen Valenz verbunden sind. Emotionale Erregung hingegen kann ebenfalls in Verbindung mit positiver Valenz auftreten.

### 2.2.2 Einfluss von Emotionen auf den Lernerfolg

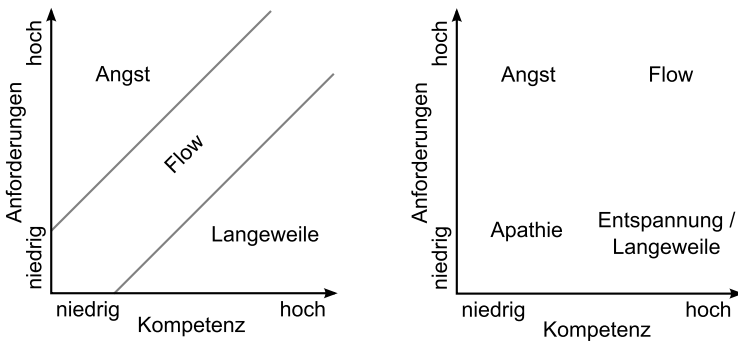
Emotionen und der Erfolg beim Lernen sind eng miteinander verbunden. So haben Emotionen einen direkten Einfluss auf die Geschwindigkeit, mit welcher Informationen verarbeitet werden (Öhman et al., 2001) und ob die Informationen überhaupt beachtet werden (z. B. Vuilleumier & Schwartz, 2001). Des Weiteren nehmen Emotionen oftmals kognitive Ressourcen in Anspruch, welche ansonsten zur Informationsverarbeitung zur Verfügung gestanden hätten (Eysenck & Calvo, 1992). Unterschiedliche Emotionen beeinflussen kognitive Prozesse auf unterschiedliche Weise: sie können beispielsweise störend, verlangsamen oder auch organisierend wirken (Pekrun et al., 2002). Hauptgegenstand dieser Arbeit ist das emotionale Erregungsniveau, welches je nach Intensität sowohl positiv als auch negativ auf die Leistungsfähigkeit einer Person bei Lernprozessen wirken kann. Dieser Zusammenhang kann mittels des in Abbildung 2.11 dargestellten Yerkes-Dodson-Gesetz (Yerkes & Dodson, 1908) veranschaulicht werden, welches den Zusammenhang zwischen der Lernperformance und der emotionalen Erregung abbildet.

Demnach hängt der Einfluss von emotionaler Erregung auf die Leistungsfähigkeit beim Lernen von der Schwierigkeit einer Aufgabe ab: Während bei einfachen Aufgaben die Performance mit zunehmender Erregung bis zu einem Sättigungsniveau ansteigt, gibt es bei schwierigeren Aufgaben ein optimales Erregungsniveau, bei dem die höchste Leistung erzielt wird. Bei zu hoher oder zu niedriger Erregung hingegen nimmt die Leistung ab. Zu den einfachen Aufgaben zählen nach Diamond et al. (2007) beispielsweise Aufgaben, die das Kurzzeitgedächtnis oder die Aufmerksamkeitsfokussierung betreffen, wohingegen Arbeitsgedächtnisaufgaben ebenso wie Multitasking oder das Treffen von Entscheidungen zu den schwierigen Aufgaben zählen. Dickman (2002) führt den Verlauf der Kurve bei schwierigen Aufgaben darauf zurück, dass niedrige Erregung, wie sie beispielsweise bei Unterforderung oder Langeweile auftritt, dazu führt, dass aufgabenrelevante Informationen nicht ausreichend aufgenommen werden. Bei zu hoher Er-



**Abbildung 2.11:** Zusammenhang von Leistung und emotionaler Erregung nach dem Yerkes-Dodson-Gesetz

regung – sei es durch Stress oder durch externe Ablenkung – fällt es insbesondere bei kognitiven Leistungen schwer, den Fokus aufrecht zu erhalten.



**Abbildung 2.12:** Das Flow-Modell nach Csikszentmihalyi. Links: ursprüngliche Version, rechts: überarbeitete Version

Eine ähnliche Annahme findet sich auch beim Flow-Modell von Csikszentmihalyi (1975). Den Kern des ursprünglichen Modells bildete die Annahme, dass Flow genau dann auftritt, wenn die Anforderungen an eine Person im Einklang mit deren Kompetenzen sind. Jedoch wurde von Csikszent-

mihalyi & Csikszentmihalyi (1988) aufgrund von Inkonsistenzen bei der Datenanalyse eine Überarbeitung des Modells dahingehend vorgenommen, welche noch heute Verwendung findet. Gemäß der überarbeiteten Version des Flow-Modells tritt Flow nur dann auf, wenn sowohl die Anforderungen als auch die Kompetenzen einer Person hoch sind. Abbildung 2.12 zeigt eine Gegenüberstellung des Originalmodells und des überarbeiteten Modells. Auch im Flow-Modell finden sich die Annahmen des Yerkes-Dodson-Gesetzes wieder, da Entspannung und Langeweile Zuständen mit einem niedrigen Erregungsniveau entsprechen, während Angst durch eine hohe emotionale Erregung gekennzeichnet ist. Folglich entspricht auch der Flow-Zustand dem Erreichen des optimalen Erregungsniveaus.

Nakamura & Csikszentmihalyi (2002) benennen folgende Faktoren als kennzeichnend für den Zustand des Flows:

- Intensive und fokussierte Konzentration auf die aktuelle Tätigkeit
- Verbindung von Handlung und Bewusstsein
- Selbstvergessenheit
- Gefühl der Kontrolle über Handlung und Umwelt
- Verlust des Zeitgefühls
- Das Gefühl, dass sich die Tätigkeit intrinsisch lohnt

Das Erreichen dieses Flow-Zustands kann somit erheblich zu einer Verbesserung des Lernergebnisses beitragen.

### 2.2.3 Das Arbeitsgedächtnis

Im Gegensatz zum Kurzzeitgedächtnis, das lediglich für die kurzfristige Speicherung von Informationen verantwortlich ist, umfasst der Begriff des Arbeitsgedächtnis die gleichzeitige Speicherung und Verarbeitung von Informationen für komplexe kognitive Aufgaben über kurze Zeit (Baddeley, 1992, 2012). Die Leistungsfähigkeit des Arbeitsgedächtnis hat einen großen Einfluss auf übergeordnete kognitive Aufgaben wie beispielsweise Leseverständnis, komplexes Lernen oder logisches Denken (Daneman & Carpenter, 1980).

Zahlreiche Studien haben gezeigt, dass die Leistungsfähigkeit des Arbeitsgedächtnisses mit zunehmendem Alter abnimmt (Hale et al., 2011;

Park & Payer, 2006). Jedoch kann ein kontinuierliches Training des Arbeitsgedächtnisses helfen, diesen Prozess zu verlangsamen. Auch im jüngeren Alter kann sich ein Training des Arbeitsgedächtnisses positiv auf die Leistungsfähigkeit in vielen Bereichen auswirken. Aus diesem Grund wurde in den letzten Jahren eine Vielzahl von Anwendungen entwickelt, die zum Training des Arbeitsgedächtnisses dienen. Auf diese wird in Abschnitt 3.4.2 eingegangen.

### 2.2.3.1 Bedeutung des Arbeitsgedächtnisses

Die Kapazität des Arbeitsgedächtnisses ist von großer Bedeutung für die kognitive Leistungsfähigkeit. Zur Veranschaulichung der Bedeutung des Arbeitsgedächtnisses für die kognitive Leistungsfähigkeit, soll zunächst der Begriff der Intelligenz erläutert werden. Es gibt verschiedene Definitionen des Intelligenzbegriffes. Eines der einflussreichsten Modelle ist die Theorie nach Cattell (1943) welche die Intelligenz in zwei Komponenten unterteilt:

*Fluide Intelligenz* beschreibt die Fähigkeit, das Denken an neue kognitive Probleme und Situationen anzupassen.

*Kristalline Intelligenz* bezieht sich auf das Faktenwissen einer Person und umfasst unter anderem Wortschatz und Rechenfähigkeit.

Im Gegensatz zur fluiden Intelligenz hängt die kristalline Intelligenz stark von kulturellen Einflüssen ab. Sie wird maßgeblich von der fluiden Intelligenz beeinflusst, da es Personen durch eine erhöhte fluide Intelligenz leichter fällt, neue Wissensbereiche zu erschließen und zu verarbeiten (Lohaus & Vierhaus, 2013).

Zwar gibt es einige Publikationen, die den Einfluss des Arbeitsgedächtnisses auf andere Aufgabengebiete kritisch betrachten, jedoch konnte in zahlreichen Forschungsarbeiten nachgewiesen werden, dass ein Training des Arbeitsgedächtnisses zu einer Steigerung der fluiden Intelligenz führt (Klingberg et al., 2002; Jaeggi et al., 2008; Borella et al., 2010; Jaeggi et al., 2010). Dies bedingt wiederum eine schnellere Verarbeitung neuer Informationen und somit eine effiziente Aneignung neuen Wissens.

Neben der Steigerung der fluiden Intelligenz gibt es auch Belege für die positive Wirkung von Arbeitsgedächtnistraining auf die Aufmerksamkeit (Thorell et al., 2009), das episodische Gedächtnis (Buschkuehl et al., 2008;

Richmond et al., 2011; Schmiedek et al., 2010; Heinzl et al., 2013) sowie auf die Verarbeitungsgeschwindigkeit (Borella et al., 2010; Heinzl et al., 2013).

### 2.2.3.2 Training des Arbeitsgedächtnisses

Zum Training des Arbeitsgedächtnisses wurde eine Reihe von Aufgaben entwickelt, die nachweislich als leistungsfördernd gelten. Ein wichtiger Unterschied zu Aufgaben zum Kurzzeitgedächtnistraining besteht darin, dass über das Speichern von Informationen hinaus auch das gleichzeitige Modifizieren von Informationen trainiert wird. Es gibt eine Reihe standardisierter Übungen zum Training des Arbeitsgedächtnisses, welche nachfolgend dargestellt sind.

#### Complex-Span-Task

Die Complex-Span-Task stellt eine Erweiterung der Simple-Span-Task dar, welche beim Training des Kurzzeitgedächtnisses zum Einsatz kommt. Bei der Simple-Span-Task wird eine Folge von verbalen oder räumlichen Objekten gezeigt, welche anschließend in der richtigen Reihenfolge reproduziert werden muss. Die Complex-Span-Task erweitert die Simple-Span-Task dahingehend, dass zwischen der Präsentation der Objekte einfache Aufgaben wie beispielsweise leichte Rechnungen bearbeitet werden müssen (Dane-man & Carpenter, 1980). Es hat sich gezeigt, dass die Complex-Span-Task zuverlässig zur Vorhersage kognitiver Fähigkeiten genutzt werden kann (Daneman & Merikle, 1996).

#### Stroop-Task

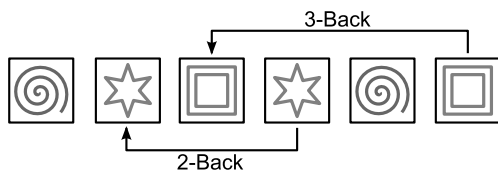
Bei der Stroop-Task besteht die Aufgabe darin, die Druckfarbe eines Farbwortes zu bestimmen, wobei Druckfarbe und Farbwort voneinander abweichen können. Die Stroop-Task geht zurück auf John Ridley Stroop, der 1935 herausfand, dass die Reaktionszeit bei der Bestimmung der Druckfarbe von Farbwörtern, bei welchen das Farbwort inkongruent zu der Schriftfarbe ist, signifikant länger dauert als die Bestimmung der Druckfarbe von Quadraten. Die Schwierigkeit bei der Stroop-Task besteht darin, dass die meisten Personen dazu neigen, ihre Aufmerksamkeit insbesondere auf das Lesen eines Wortes statt auf die Druckfarbe zu richten (MacLeod,



1991). Diese Aufgabe spiegelt daher insbesondere die Fähigkeit wider, die Aufmerksamkeit auf bestimmte Dinge zu richten (Shipstead et al., 2012).

### N-Back-Task

Bei der sogenannten  $n$ -Back-Task (Kirchner, 1958) wird den Trainierenden eine Liste von Objekten gezeigt. Die Aufgabe besteht darin, zu bestimmen, ob das aktuelle Objekt mit dem Objekt, das  $n$  Schritte zuvor präsentiert wurde, übereinstimmt. So müssen bei der 2-Back-Task die letzten beiden Objekte gemerkt und darauf basierend bestimmt werden, ob das aktuelle Objekt mit dem vorletzten Objekt übereinstimmt. Üblicherweise werden hierbei visuelle oder akustische Objekte verwendet. Abbildung 2.13 veranschaulicht die Vorgehensweise bei dieser Aufgabe.



**Abbildung 2.13:** Ablauf der  $n$ -Back-Task

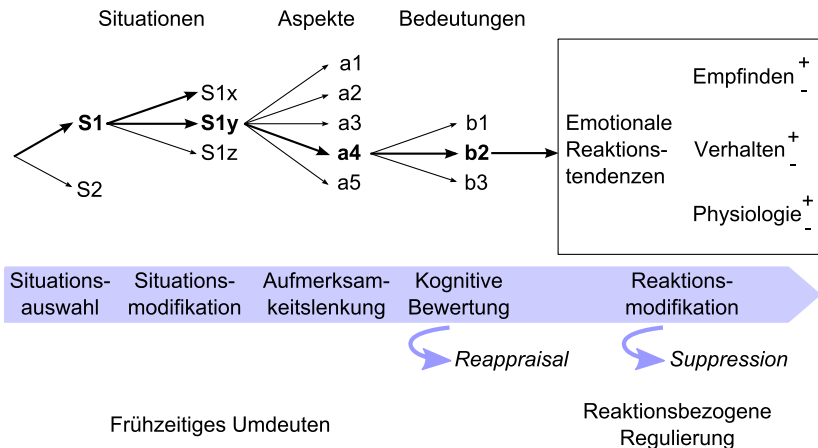
Eine Erweiterung der  $n$ -Back-Task stellt die von Jaeggi et al. (2003) eingeführte Dual  $n$ -Back-Task dar. Bei dieser Variante werden zwei unterschiedliche Objekte gleichzeitig präsentiert – oft auf unterschiedliche Weise wie beispielsweise akustisch und visuell. Bei der  $n$ -Back-Task handelt es sich – nicht zuletzt aufgrund der sehr flexibel zu variierenden Schwierigkeit – um eine der am häufigsten verwendeten Aufgaben zum Arbeitsgedächtnistraining (Conway et al., 2005).

### Running Memory-Span Task

Beim Running Memory-Span (Pollack et al., 1959) wird dem Trainierenden eine Liste von Objekten präsentiert. Zu einem zufälligen Zeitpunkt wird die Präsentation unterbrochen und die zuletzt präsentierten Objekte müssen wiedergegeben werden.

### 2.2.4 Emotionsregulation

Emotionen beeinflussen uns in vielen Aspekten des täglichen Lebens, und auch beim Lernen gibt es emotionale Zustände, die für den Lernerfolg förderlicher sind, als andere. Die Fähigkeit, seine Emotionen entsprechend den Anforderungen einer bestimmten Situation zu regulieren, ist daher von großer Bedeutung. Das am weitesten verbreitete Emotionsregulationsmodell ist das Prozessmodell von Gross (1998), bei welchem von verschiedenen Phasen der Emotionsgenerierung ausgegangen wird. Die einzelnen Phasen der Emotionsgenerierung in diesem Prozessmodell sowie die damit verbundenen Emotionsregulationsmöglichkeiten sind in Abbildung 2.14 dargestellt. Gross unterscheidet in diesem Modell zwischen frühzeitigem Umdeuten („*antecedent-focused emotion regulation*“), welches bereits vor der Entstehung einer Emotion stattfindet, sowie reaktionsbezogener Regulierung („*response-focused emotion regulation*“) einer emotionsauslösenden Situation, welche erst erfolgt, nachdem die emotionale Reaktion bereits hervorgerufen wurde.



**Abbildung 2.14:** Prozessmodell der Emotionsregulation nach John & Gross (2004)

Der Regulationsprozess beginnt mit der Situationsauswahl. So werden bevorstehende unangenehme Situationen häufig vermieden, weil diese un-

gewünschte Emotionen hervorrufen können (z. B. Aspinwall & Taylor, 1997; Zeelenberg & Pieters, 2007). Im Rahmen der Situationsmodifikation kann diese anschließend so angepasst werden, dass sich die emotionale Auswirkung der Situation verändert. Im nächsten Schritt kann die Aufmerksamkeit auf bestimmte Aspekte einer Situation gelenkt werden, die in dem jeweiligen Moment relevant sind. Ist dies geschehen, kann eine kognitive Neubewertung der gewählten Aspekte erfolgen. Die letzte Phase der Emotionsregulation erfolgt gewissermaßen als Reaktion auf die hervorgerufene Emotion, indem die Auswirkungen dieser Emotion unterdrückt werden.

Im Rahmen dieser Arbeit werden insbesondere die Strategie der kognitiven Neubewertung einer Situation („*Reappraisal*“) sowie die Unterdrückung der durch eine Emotion ausgelösten Reaktion („*Suppression*“) betrachtet, da diese nach John & Gross (2004) im täglichen Leben sehr häufig verwendet werden.

## 2.3 Mathematische und algorithmische Grundlagen

Im Folgenden wird ein Überblick über die statistischen Methoden gegeben, die im Rahmen dieser Arbeit verwendet wurden. Soweit nicht anders gekennzeichnet, sind die Beschreibungen Bortz & Schuster (2010) entnommen. Für weitergehende Informationen über die statistische Analyse von Daten sei auf Bortz & Schuster (2010) verwiesen.

### 2.3.1 Feature-Extraktion aus physiologischen Signalen

Für die Klassifikation von Emotionen basierend auf physiologischen Parametern müssen zunächst Features aus den Rohsignalen extrahiert werden. Im Rahmen dieser Arbeit wurden eine Reihe standardisierter Verfahren verwendet, welche nachfolgend vorgestellt werden. Eine Beschreibung der weiteren Verfahren, welche speziell für die im Rahmen dieser Arbeit durchgeführten Studien entwickelt wurden, findet sich in den jeweiligen Kapiteln der einzelnen Studien.

### 2.3.1.1 Kardiovaskuläre Daten

Im Rahmen dieser Arbeit werden kardiovaskuläre Daten sowohl mittels PPG- als auch mit EKG-Sensoren aufgezeichnet. Die Vorgehensweise bei der Analyse dieser Daten wird im folgenden beschrieben.

#### Feature-Extraktion aus dem EKG-Signal

Wie in Abschnitt 2.1.2.2 erläutert, wird zur Bestimmung des Zeitpunktes einzelner Herzschläge in der Regel die R-Zacke des EKG-Signals herangezogen, da diese aufgrund ihrer spitzen Form sehr genau die Bestimmung des exakten Zeitpunktes eines Herzschlags zulässt.

Zur Bestimmung der R-Zacken wird zunächst mittels des von Hamilton (2002) entwickelten Open Source ECG Analysis Algorithm (OSEA) die Lage der QRS-Komplexe bestimmt. Der Vorteil dieses Algorithmus liegt darin, dass er mit geringem Rechenaufwand sehr gute Erkennungsleistungen erzielt. Im Vergleich zu anderen Algorithmen zeigte sich, dass der OSEA-Algorithmus insbesondere bei artefaktbehafteten Signalen über die größte Sensitivität und positive Prädiktivität verfügt (Ottenbacher et al., 2008). Nach der Detektion der QRS-Komplexe kann der exakte Zeitpunkt der R-Zacke als Maximum über die benachbarten Samples ermittelt werden.

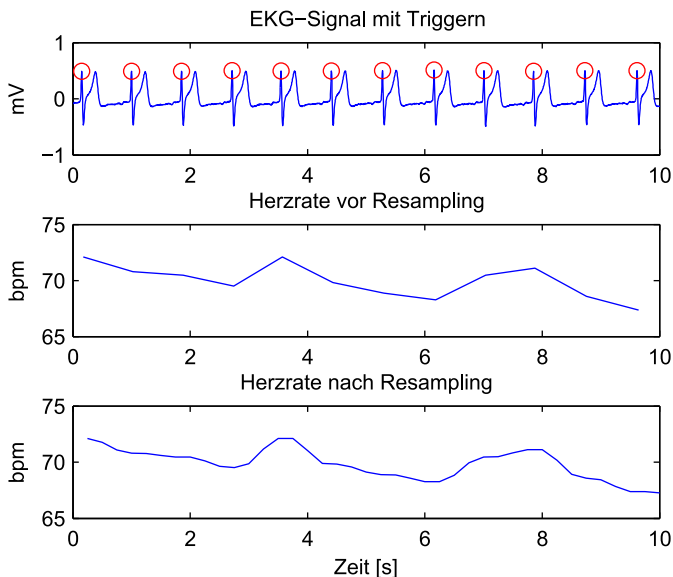
Zur Validierung der detektierten R-Zacken wird im Rahmen dieser Arbeit das in Clifford (2007) beschriebene schwellwertbasierte Verfahren verwendet. Die als valide eingestuft R-Zacken werden gemäß der Richtlinien des WFDB Programmer's Guide (Moody, 2013) mit  $N$  (*normal beat*) markiert, während Artefakte mit  $Q$  (*unclassifiable beat*) annotiert werden. Die Intervalle zwischen zwei validen R-Zacken werden im Folgenden als  $NN$ -Intervalle bezeichnet.

Basierend auf den  $NN$ -Intervallen kann die Herzrate als Reziproke der  $NN$ -Intervalle berechnet werden:

$$HR_n = 60000/RR_n \quad (2.3)$$

Hieraus ergibt sich zunächst ein ungleichmäßig abgetastetes Signal. Zur weiteren Analyse der Herzrate empfiehlt sich daher ein Resampling der Daten. Der Vorteil eines Resamplings liegt darin, dass das neue Signal den tatsächlichen Verlauf der Herzrate besser widerspiegelt. Insbesondere bei der Bildung von Mittelwerten über einen bestimmten Zeitraum werden

bei Analysen der Herzrate, bei denen kein Resampling durchgeführt wird, schnellere Schläge über- und langsamere Schläge unterschätzt. Im Rahmen dieser Arbeit wird das Resampling mittels des von Berger et al. (1989) entwickelten Verfahrens vorgenommen. Gemäß den Empfehlungen von Singh et al. (2004) wird hierfür eine Samplingrate von 8 Hz verwendet. Abbildung 2.15 veranschaulicht die Schritte von der R-Zacken-Detektion



**Abbildung 2.15:** Ermittlung der Herzrate aus dem EKG-Signal. Die roten Markierungen kennzeichnen die detektierten R-Zacken

zum resampelten Signal.

Neben der Herzrate wird als Maß für die Variabilität der Herzrate oftmals die Herzratenvariabilität (HRV) herangezogen. Die Analyse der Herzratenvariabilität kann auf zwei unterschiedliche Weisen erfolgen: Zum einen können Features im Zeitbereich berechnet werden, auf der anderen Seite ist auch eine Analyse von Features im Frequenzbereich möglich. Bei der Analyse im Zeitbereich kann zwischen statistischen und geometrischen Maßen sowie der Analyse der Poincaré-Geometrie unterschieden werden.

Folgende statistische Parameter werden im Rahmen dieser Arbeit be-

rechnet:

- *meanNN*: Mittelwert der NN-Intervalle in einem Zeitfenster (entspricht dem Erwartungswert  $\mu$ )

$$\text{meanNN} = \mu = \frac{1}{n} \sum_{i=1}^n NN_i \quad (2.4)$$

- *SDNN*: Standardabweichung der NN-Intervalle

$$\text{SDNN} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (NN_i - \mu)^2} \quad (2.5)$$

- *RMSSD*: Variation aufeinander folgender NN-Intervalle

$$\text{RMSSD} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n-1} (NN_{i+1} - NN_i)^2} \quad (2.6)$$

- *pNNx* Anzahl benachbarter NN-Intervalle, zwischen denen die Differenz mehr als  $x$  ms beträgt dividiert durch die Anzahl aller NN-Intervalle in dem dazugehörigen Zeitfenster.

Der gebräuchliche Wert für *pNNx* liegt bei 50 ms. Jedoch werden in Anlehnung an Mietus et al. (2002) zusätzlich *pNN12* sowie *pNN50* berechnet, da sich gezeigt hat, dass diese oftmals einen genaueren Aufschluss über die Veränderungen der HRV zulassen.

Die Beschreibung der Poincaré-Plot-Geometrie erfolgt durch die Standardabweichungen der NN-Intervalle zu den Diagonalen  $y = x$  sowie  $y = -x + 2 \cdot \text{meanNN}$ . Die Standardabweichungen werden als SD1 und SD2 bezeichnet. Zusätzlich kann das Verhältnis der Standardabweichungen SD1/SD2 analysiert werden (Tulppo et al., 1996).

Bei geometrischen Maßen wird die Verteilung der RR-Intervalle in einem Histogramm betrachtet. Im Gegensatz zu anderen Features im Zeitbereich sind hierbei jedoch verhältnismäßig lange Analysezeiträume von mindestens 20 Minuten erforderlich, um valide Ergebnisse zu erhalten (Malik, 1996).

Zur Transformation vom Zeit- in den Frequenzbereich können verschiedene Methoden eingesetzt werden. Im Rahmen dieser Arbeit wird die Lomb-Transformation (Lomb, 1976; Scargle, 1982) verwendet, da sich diese im Gegensatz zu anderen Methoden dadurch auszeichnet, dass auch ungleichmäßig abgetastete Daten ohne vorheriges Resampling verarbeitet werden können.

Im Frequenzbereich werden bei der HRV-Analyse verschiedene Frequenzbänder unterschieden (Malik, 1996):

Parameter	Beschreibung	Frequenzbereich
ULF	ultra low frequency	< 0,003 Hz
VLF	very low frequency	0,003 Hz bis 0,04 Hz
LF	low frequency	0,04 Hz bis 0,15 Hz
HF	high frequency	0,15 Hz bis 0,40 Hz

**Tabelle 2.1:** Frequenzbänder für die HRV-Analyse (Malik, 1996)

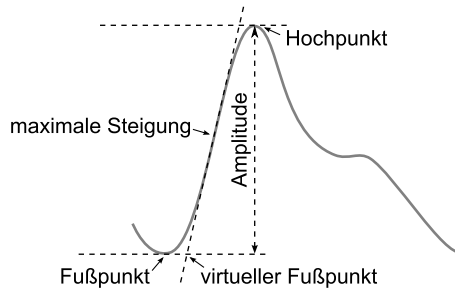
In dieser Arbeit werden sowohl das LF- als auch das HF-Frequenzspektrum analysiert. VLF und ULF werden nicht betrachtet, da die entsprechenden Frequenzbänder bei Zeitfenstern < 5 Minuten nicht vollständig aufgelöst werden können. Aus den Kennwerten für LF und HF kann zudem das LF/HF-Ratio ermittelt werden.

### Feature-Extraktion aus dem PPG-Signal

Die Pulswelle ist durch eine Reihe charakteristischer Merkmale gekennzeichnet:

- *Hochpunkt:* Maximum der Pulskurve
- *Maximale Steigung:* Maximaler Anstieg der Pulskurve
- *Virtueller Fußpunkt:* Schnittpunkt der Tangente an der maximalen Steigung mit dem Fußpunkt
- *Fußpunkt:* Minimum der Pulskurve

Die Amplitudenhöhe ergibt sich als Differenz zwischen Hochpunkt und Tiefpunkt der Pulswelle.



**Abbildung 2.16:** Charakteristische Merkmale der Puls-PPG-Kurve

### Kombination von EKG- und PPG-Signal

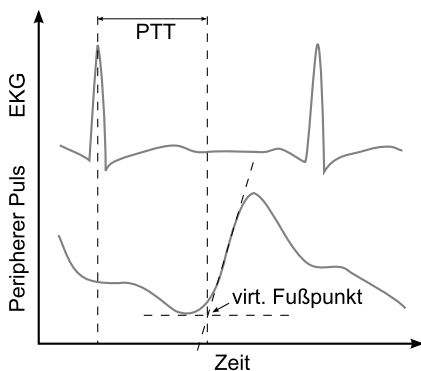
Stehen sowohl ein EKG- als auch ein PPG-Signal zeitsynchron zur Verfügung, ermöglicht die Kombination beider Signale die Bestimmung der PTT. Aus beiden Signalen müssen hierzu eindeutige Merkmale extrahiert werden. Bei dem EKG-Signal wird üblicherweise die R-Zacke verwendet. Bei der PPG-Kurve gibt es wie in Abbildung 2.16 dargestellt eine Reihe von möglichen Merkmalen, jedoch können Merkmale, die nach dem systolischen Anstieg liegen von den Reflexionen der Puls-PPG-Kurve in den Gefäßen verfälscht werden. Aus diesem Grund hat sich der virtuelle Fußpunkt als robustes Merkmal erwiesen, da er Informationen über den Fußpunkt sowie die maximale Steigung der Puls-PPG-Kurve enthält (Elter, 2001). Abbildung 2.17 veranschaulicht die Berechnung der PTT aus dem EKG- und dem PPG-Signal.

Aufgrund der physiologischen und anatomischen Randbedingungen gelten bei einer PPG-Messung am Finger folgende Kriterien für valide Werte der PTT:

- Maximum der Puls-PPG-Kurve mindestens 110 und maximal 475 ms von der vorhergehenden R-Zacke entfernt
- Differenz zwischen dem zu der vorhergehenden Puls-PPG-Kurve ermittelten Interbeat-Intervall und dem dazugehörigen NN-Intervall des EKG-Signals unter 40 ms

Anhand dieser Kriterien können artefaktbehaftete Signalteile leicht identifiziert und von der weiteren Analyse ausgeschlossen werden.





**Abbildung 2.17:** Berechnung der PTT als Differenz zwischen R-Zacke und dem virtuellen Fußpunkt

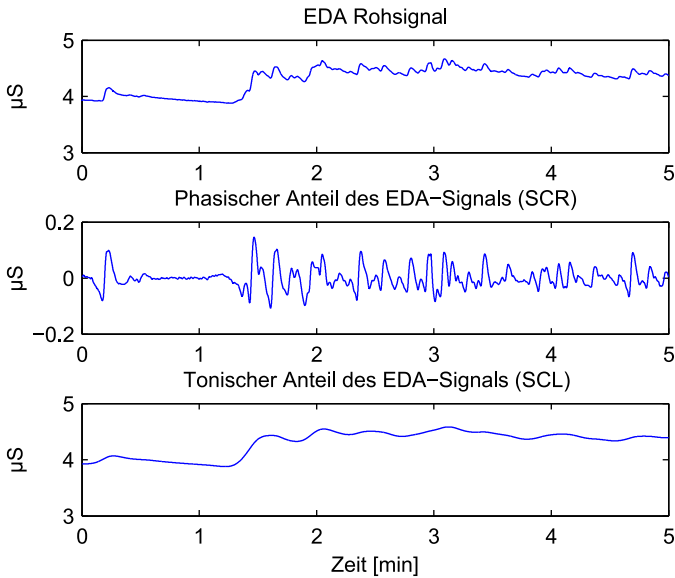
### 2.3.1.2 Elektrodermale Aktivität

EDA setzt sich aus zwei Komponenten zusammen: dem sich langsam verändernden tonischen Hautleitwertsniveau (auch *Skin conductance level*, SCL)) sowie den phasischen Hautleitwertsreaktionen (*Skin conductance responses*, SCR)). Letztere gehen oft mit einem konkreten Stimulus einher.

Bei der Analyse von EDA können sowohl tonische als auch phasische Veränderungen betrachtet werden. In Abbildung 2.18 ist die Ableitung der phasischen und tonischen Aktivität aus dem EDA-Signal dargestellt.

Zusätzlich wird unterschieden zwischen SCRs, die als Reaktion auf einen bestimmten Stimulus auftreten, sowie den so genannten Spontanfluktuationen (*non-significant SCR*, NS-SCR), die ohne zuordenbaren Reiz vorkommen. Von letzteren wird oftmals auch die Häufigkeit des Auftretens in einem bestimmten Zeitfenster analysiert. (Dawson et al., 2007)

Bei den auf dem EDA-Signal berechneten Features kann unterschieden werden zwischen event- und zeitbasierten Features. Die eventbasierten Features basieren auf den detektierten phasischen Reaktionen und können nur dann berechnet werden, wenn tatsächlich eine elektrodermale Reaktion auftritt. Die zeitbasierten Features hingegen können direkt aus dem kontinuierlichen Signal abgeleitet werden.



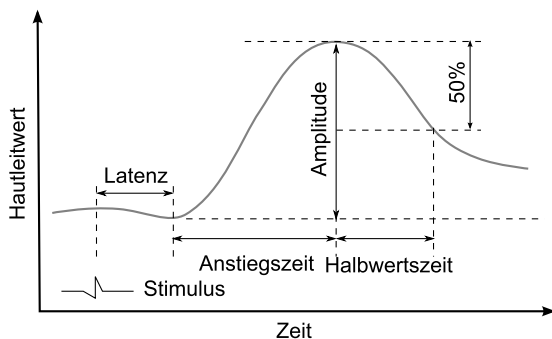
**Abbildung 2.18:** Trennung des EDA-Signals in phasische und tonische Aktivität

### Eventbasierte Features

Für die Berechnung der eventbasierten Features müssen zunächst die einzelnen phasischen Hautleitwertsreaktionen detektiert werden. Abbildung 2.19 zeigt die typischen Parameter, welche üblicherweise analysiert werden.

Für die Detektion der Reaktionen wird das von Zimmermann (2012) entwickelte Verfahren verwendet, welches speziell für echtzeitnahe Anwendungen entworfen wurde. Es gibt eine Reihe weiterer etablierter Verfahren wie beispielsweise das faltungsbasierte Verfahren nach Benedek & Kaernbach (2010). Jedoch ist dieses bei ähnlicher Genauigkeit weitaus laufzeitintensiver.

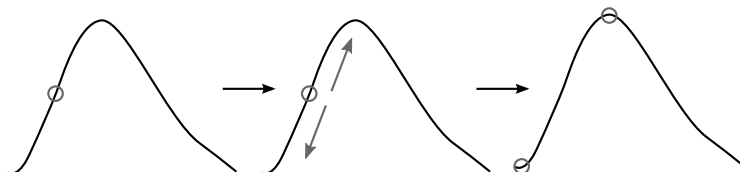
Die Detektion der EDA-Reaktionen erfolgt in drei Schritten: Zunächst wird nach Wendepunkten des vorgefilterten EDA-Signals mit ansteigender Flanke gesucht. Ausgehend von diesen Wendepunkten wird im nächsten Schritt der jeweilige vorhergehende Tiefpunkt sowie der nachfolgende



**Abbildung 2.19:** Typische Parameter elektrodermalen Reaktionen

Hochpunkt innerhalb eines Zeitfensters von 3 Sekunden bestimmt. Wird innerhalb dieses Zeitfensters kein Tief- bzw. Hochpunkt gefunden, wird der nächste Wendepunkt betrachtet. Die einzelnen Detektionsschritte sind in Abbildung 2.20 schematisch dargestellt. Da die Berechnung der Amplitudenhöhe durch überlagerte Reaktionen verfälscht werden kann, findet im letzten Schritt eine Fallunterscheidung zur Erkennung überlagerter Reaktionen statt. Im Falle von überlagerten Reaktionen wird zur weiteren Berechnung der Features der theoretische Verlauf der überlagerten Reaktionen mittels einer Exponentialfunktion bestimmt.

Die Amplitude einer Reaktion wird bestimmt als die Höhendifferenz zwischen einem (theoretischen) Tiefpunkt und dem Hochpunkt der jeweiligen Reaktion. Die Anstiegszeit gibt die Zeit an, die zwischen Tiefpunkt



**Abbildung 2.20:** Vorgehensweise bei der Detektion elektrodermalen Reaktionen nach Zimmermann (2012)

und Hochpunkt liegt. Halbwertszeit wird gemessen als die Zeit zwischen Hochpunkt und dem Absinken der (theoretischen) Reaktionen auf 50 % der Amplitudenhöhe. Zusätzlich kann die Signalenergie als Produkt aus der Amplitude und der Anstiegszeit berechnet werden.

Im Anschluss an die Detektion der SCRs wird eine Plausibilitätskontrolle anhand der in Dawson et al. (2007) empfohlenen Schwellwerte durchgeführt. SCRs, welche außerhalb dieser Schwellwerte liegen, werden verworfen.

### Zeitbasierte Features

Das wohl prominenteste der in dieser Arbeit betrachteten zeitbasierten Features stellt das Hautleitwertsniveau dar. Für die Extraktion der tonischen Veränderungen des EDA-Signals wird in Anlehnung an Dawson et al. (2007) eine Tiefpassfilterung mit einem Butterworth-Filter 2. Ordnung und einer Grenzfrequenz von 0,1 Hz vorgenommen. Auf diese Weise werden alle phasischen Anteile aus dem Signal eliminiert.

Des Weiteren werden in Anlehnung an Kim & André (2008) Anzahl und Amplitude der Nulldurchgänge untersucht. Hierzu werden die Daten zunächst mit einem Butterworth-Hochpassfilter mit einer Grenzfrequenz von 0,2 Hz gefiltert. Im nächsten Schritt erfolgt eine Bestimmung der Nulldurchgänge des gefilterten Signals. Zwischen zwei benachbarten Nulldurchgängen wird jeweils das Maximum bzw. Minimum des Signals bestimmt. Die Amplitude des Nulldurchgangs ergibt sich als Differenz zwischen einem benachbartem Minimum und dem darauf folgendem Maximum. Für die Ermittlung der Anzahl der Nulldurchgänge wird die minimal zu berücksichtigende Amplitudenhöhe auf 0,1  $\mu\text{S}$  festgelegt. Für die Analyse der Amplitudenhöhe selbst muss diese mindestens 0,01  $\mu\text{S}$  betragen.

In Anlehnung an Fleureau et al. (2013) wird zusätzlich die positive Ableitung des EDA-Signals bestimmt. Hierfür wird die erste Ableitung des EDA-Signals gebildet und im nächsten Schritt alle negativen Werte verworfen. Somit beinhaltet das berechnete Signal nur noch Informationen über die sympathisch ausgelösten phasischen Veränderungen.

Darüber hinaus werden in dieser Arbeit die sogenannten Hjorth-Parameter (Hjorth, 1970) zur Analyse der Veränderungen des EDA-Signals herangezogen. Diese ursprünglich aus der EEG-Analyse stammenden Parameter können wie folgt berechnet werden:

$$\text{Aktivität} = \text{var}(x(t)) \quad (2.7)$$

$$\text{Mobilität} = \sqrt{\frac{\text{Aktivität} \left( \frac{dx(t)}{dt} \right)}{\text{Aktivität}(x(t))}} \quad (2.8)$$

$$\text{Komplexität} = \sqrt{\frac{\text{Mobilität} \left( \frac{dx(t)}{dt} \right)}{\text{Mobilität}(x(t))}} \quad (2.9)$$

Die Aktivität gibt hierbei Informationen über die Signalenergie, die Mobilität spiegelt die mittlere Frequenz wider und die Komplexität erlaubt Rückschlüsse auf die Veränderung der Frequenz des Signals.

## 2.3.2 Maschinelles Lernen und Klassifikation

Unter maschinellem Lernen versteht man die Fähigkeit eines Programmes, aus Erfahrung zu lernen. Dies bedeutet, dass sich die Ausgabe eines Programmes basierend auf vorher gesammelten Informationen verändern kann. Nachfolgend sind die Klassifikationsverfahren aus dem Bereich des maschinellen Lernens vorgestellt, welche im Rahmen dieser Arbeit verwendet werden.

### 2.3.2.1 Diskriminanzanalyse

Ziel der Diskriminanzanalyse ist das Finden einer optimalen Kombination der vorhandenen Featurevektoren zur Trennung verschiedener Klassen. Diese Kombination kann entweder – im Falle der Linearen Diskriminanzanalyse (LDA) – linear, oder aber quadratisch sein wie bei der quadratischen Diskriminanzanalyse (QDA). Die Feature-Kombination kann als Hyperebene in einem  $n$ -dimensionalen Raum gesehen werden, welche den Raum in die unterschiedlichen Klassen aufteilt, wobei  $n$  die Dimensionalität des Featurevektors darstellt.

Das Grundprinzip der Diskriminanzanalyse sei hier an einem 2-Klassenproblem mit den Klassen  $C_1$  und  $C_2$  erläutert. Die Diskriminanzfunktion  $D(\mathbf{x})$  nimmt hierbei für einen Featurevektor  $\mathbf{x}$  folgende Form an:

$$D(\mathbf{x}) = \mathbf{w}\mathbf{x} + b \quad (2.10)$$

wobei  $\mathbf{w}$  die Gewichtungskoeffizienten darstellt und  $b$  den Achsenabschnitt der Funktion.

Die Gewichtungskoeffizienten der Diskriminanzfunktion werden hierbei so gewählt, dass das Verhältnis der Varianz zwischen den Klassen zu der Varianz innerhalb der Klassen maximiert wird.

Liegt das Ergebnis der Diskriminanzfunktion für einen Featurevektor über einem vorgegebenen Schwellwert  $T$ , so wird dieser Featurevektor der Klasse  $C_1$  zugeordnet. Nimmt die Diskriminanzfunktion einen Wert unter diesem Schwellwert an, so erfolgt eine Zuordnung zu Klasse  $C_2$ .

### 2.3.2.2 k-Nearest-Neighbour

Bei dem k-Nearest-Neighbour-Algorithmus (kNN) werden zur Klassifikation eines Featurevektors jeweils die  $k$  nahegelegendsten Featurevektoren betrachtet. Ein zu klassifizierender Datenpunkt wird der Klasse zugeordnet, zu welcher das Distanzmaß der  $k$  benachbarten Elemente am geringsten ist.

Als Distanzmaß wird zumeist der euklidische Abstand herangezogen. Für einen Featurevektor  $\mathbf{r}$  kann der Abstand  $d$  zu einem benachbarten Featurevektor  $\mathbf{r}_i$  somit wie folgt bestimmt werden:

$$d(\mathbf{r}, \mathbf{r}_i) = \sqrt{\left( \sum_{n=1}^N |r_n - r_{i,n}|^2 \right)} \quad (2.11)$$

Der Vorteil des kNN-Algorithmus kann darin gesehen werden, dass sich durch dieses Verfahren sehr komplexe Entscheidungsgebiete gewinnen lassen (Kroschel et al., 2011).

### 2.3.2.3 Entscheidungsbäume

Entscheidungsbäume stellen eine weitere Möglichkeit zur Klassifikation von Daten dar. Bei Entscheidungsbäumen handelt es sich um eine spezielle Form der Graphen. Ein Baum ist somit definiert aus einer Menge von Knoten und Kanten mit der Bedingung, dass es keine kreisförmigen

Verbindungen gibt und der Graph zusammenhängend aufgebaut sein muss. Der Graph hat genau einen Knoten ohne eingehende Kanten, welcher als Wurzel des Entscheidungsbaumes bezeichnet wird. Knoten ohne ausgehende Kanten werden als Blätter bezeichnet. Bei der Klassifikation eines Featurevektors werden die einzelnen Knoten des Entscheidungsbaums von der Wurzel startend bis hin zu den Blättern durchlaufen. Die Blätter enthalten Informationen über die Zuordnung des Featurevektors zu einer Klasse (Kroschel et al., 2011).

Beim Aufbau eines Entscheidungsbaumes gibt es verschiedene Methoden. Im Rahmen dieser Arbeit werden binäre Entscheidungsbäume nach folgendem Verfahren aufgebaut:

1. Finde das Kriterium das die Menge der Eingabe-Featurevektoren am besten trennt
2. Erstelle den Split
3. Wiederhole diese Vorgehensweise rekursiv für jeden der Kind-Knoten

Das Training endet, wenn alle Knoten nur noch Featurevektoren einer einzigen Klasse enthalten.

### 2.3.2.4 Support Vector Machines

Das Grundprinzip von Support Vector Machines (SVM) besteht darin, dass in einem  $n$ -dimensionalen Raum, der Daten aus verschiedenen Klassen enthält, eine Hyperebene zur Trennung der Klassen eingefügt wird. Diese Hyperebene wird derart eingefügt, dass die Differenz zwischen Daten aus verschiedenen Klassen maximiert wird.

Für ein Trainingsset mit Instanz-Label-Paaren  $(x_i, y_i), i = 1, \dots, l$  mit  $x_i \in R^n$  und  $y \in \{1, -1\}^l$  gilt es das folgende Optimierungsproblem zu lösen:

$$\min_{w, b, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \tag{2.12}$$

mit der Nebenbedingung  $y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i$

$$\xi_i \geq 0$$

Die Schlupfvariable  $\xi_i$  gibt hierbei den Grad der Fehlklassifikationen von  $x_i$  an,  $C > 0$  ist der Strafparameter des Fehlers.  $w$  ist der Normalenvektor der trennenden Hyperebene. Die Funktion  $\phi$  bildet den Trainingsvektor in einen höherdimensionalen Raum ab. In diesem höherdimensionalen Raum findet die SVM eine linear trennende Hyperebene mit maximalem Abstand.

Es gibt unterschiedliche Kernel-Funktionen  $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ , die für die Klassifikation mittels SVM verwendet werden können:

- Lineare Kernels:

$$K(x_i, x_j) = x_i^T x_j \quad (2.13)$$

- Radial basis function (RBF) Kernels:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (2.14)$$

( $\gamma$  ist der Kernel-Parameter, der die RBF-Weite definiert)

Bei Verwendung des linearen Kernels entspricht der Feature-Raum dem Eingaberaum. Im Gegensatz hierzu werden die Samples bei Verwendung eines RBF-Kernels nichtlinear in einen höherdimensionalen Raum transformiert. Dies bedeutet, dass RBF-Kernels auch mit nichtlinearen Beziehungen zwischen Klassen und Attributen umgehen können.

Ursprünglich wurden SVMs für binäre Klassifikationsprobleme entwickelt. Für die Multiklassen-Klassifikation kann der sogenannte *one-against-one*-Ansatz verwendet werden, bei welchem  $k \cdot (k - 1)/2$  Klassifikatoren verwendet werden. Jeder Klassifikator trainiert Daten aus zwei verschiedenen Klassen entsprechend des in Gleichung 2.12 dargestellten Klassifikationsproblems.

Für weitere Details zur Klassifikation mit SVMs sei auf Burges (1998) verwiesen.

### 2.3.2.5 Bewertung von Klassifikatoren

Für die Beurteilung der Güte von Klassifikatoren gibt es verschiedene Maße, welche nachfolgend dargestellt werden.

Sei  $M$  die Menge der Klassen und  $\mathcal{T}$  eine Teststichprobe mit  $|\mathcal{T}|$  Elementen, welche genau einer Klasse  $i \in \{1, \dots, M\}$  zugeordnet werden können.



Die Entscheidung für die Zuordnung eines Elements zu einer Klasse kann als Abbildung  $d : \mathbf{R} \rightarrow \{1, \dots, M\}$  betrachtet werden.

Für ein Zwei-Klassenproblem können die klassifizierten Elemente je nach Klassifikationsergebnis unterschiedlichen Kategorien zugeordnet werden. Diese sind in nachfolgender Tabelle dargestellt.

	Referenz richtig	Referenz falsch
Hypothese richtig	true positive (TP)	false positive (FP)
Hypothese falsch	false negative (FN)	true negative (TN)

**Tabelle 2.2:** Kontingenzmatrix für Klassifikationsergebnisse

Für eine Klasse  $i \in M$  können diese Maße wie folgt definiert werden:

$$\text{TP}_i = \{ \mathbf{r} \in \mathcal{T} | d(\mathbf{r}) = i \}, \quad (2.15)$$

$$\text{FP}_i = \{ \mathbf{r} \in \mathcal{T} - \mathcal{T}_i | d(\mathbf{r}) = i \}, \quad (2.16)$$

$$\text{TN}_i = \{ \mathbf{r} \in \mathcal{T} - \mathcal{T}_i | d(\mathbf{r}) \neq i \}, \quad (2.17)$$

$$\text{FN}_i = \{ \mathbf{r} \in \mathcal{T} | d(\mathbf{r}) \neq i \}. \quad (2.18)$$

Das wohl gebräuchlichste Maß für die Bewertung von Klassifikatoren stellt die Klassifikationsgenauigkeit (engl. *Recall*) dar. Diese berechnet sich als Quotient aus der Anzahl richtig erkannter Muster und dem Gesamtumfang der Teststichprobe:

$$\text{RE} = \frac{\sum_{i=1}^M |\{ \mathbf{r} \in \mathcal{T} | d(\mathbf{r}) = i \}|}{|\mathcal{T}|}. \quad (2.19)$$

Zusätzlich zu der Gesamtgenauigkeit lässt sich auch die klasenspezifische Erkennungsrate definieren, die oftmals auch als *Sensitivität* bezeichnet wird:

$$\text{RE}_i = \frac{|\{ \mathbf{r} \in \mathcal{T} | d(\mathbf{r}) = i \}|}{|\mathcal{T}_i|} = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i}. \quad (2.20)$$

Mit der A-priori-Klassenwahrscheinlichkeit  $p_i = |\mathcal{T}_i|/|\mathcal{T}|$  für eine Klasse  $M_i$  ergibt sich für die Genauigkeit folgender Zusammenhang:

$$\text{RE} = \sum_{i=1}^M p_i \text{RE}_i. \quad (2.21)$$

Aufgrund der Gewichtung mit der A-priori-Klassenwahrscheinlichkeit wird die Klassifikationsgenauigkeit oftmals auch als *Weighted-Average-Recall* (WAR) bezeichnet. Insbesondere bei unbalancierten Stichproben ist dieses Maß zur Bewertung der Unterscheidungsfähigkeit eines Klassifikators nur bedingt geeignet. Für derartige Fälle empfiehlt sich die Verwendung des *Unweighted-Average-Recall* (UAR):

$$\text{UAR} = \frac{\sum_{i=1}^M \text{RE}_i}{M}. \quad (2.22)$$

Ein weiteres Gütemaß für die Bewertung eines Klassifikators stellt die *Präzision* oder *positive Prädiktivität* dar:

$$\text{PR}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}. \quad (2.23)$$

Wird im Rahmen der Optimierung eines Klassifikators die Sensitivität optimiert, hat dies in der Regel ein Sinken der positiven Prädiktivität zur Folge. Aus diesem Grund steht mit dem *F-Maß* ein Gütemaß zur Verfügung, welches beide Kennwerte durch Bildung des harmonischen Mittels vereint:

$$F_i = 2 \cdot \frac{\text{RE}_i \text{PR}_i}{\text{RE}_i + \text{PR}_i} \quad (2.24)$$

Analog zu Sensitivität und positiver Prädiktivität können zusätzlich die *Spezifität*  $\text{SP}_i$  sowie der *negative Vorhersagewert*  $\text{NV}_i$  definiert werden:

$$\text{SP}_i = \frac{\text{TN}_i}{\text{TN}_i + \text{FP}_i}, \quad (2.25)$$

$$\text{NV}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}. \quad (2.26)$$

Bei der Klassifikation von Events wie beispielsweise dem Auftreten von R-Zacken ist die Zahl der true negatives (TN) nicht definiert. Aus diesem Grund kann in derartigen Fällen keine Berechnung der Spezifität erfolgen.



## 3 Stand der Technik

Seit dem ersten Aufkommen des Begriffes des *Affective Computing* im Jahr 1995 ist das Interesse an der Integration von Emotionen in die Mensch-Maschine-Interaktion stark gestiegen. Dieses Kapitel dient dazu, einen Überblick über die Möglichkeiten der Messung von Emotionen sowie den Bereich des Affective Computing zu geben. Abschließend werden affektive Lerntechnologien sowie die gängigsten Anwendungen zum Training des Arbeitsgedächtnisses vorgestellt.

### 3.1 Systeme zur Aufzeichnung physiologischer Daten

Für die Aufnahme physiologischer Daten steht eine Vielzahl verschiedener Systeme zur Verfügung, welche sich in zahlreichen Aspekten unterscheiden. Diese Unterscheidungen sind sowohl bei den funktionalen Eigenschaften, als auch bei den Einsatzszenarien, für die die Geräte entwickelt wurden, deutlich erkennbar. Führt man sich vor Augen, dass die Einsatzbereiche physiologischer Messgeräte von der simplen Berechnung eines Pulswertes beim Sport bis hin zu komplexen medizinischen Analysen reichen, erklärt dies die große Menge an verfügbaren Geräten.

In der vorliegenden Arbeit wurden sowohl Studien im Labor als auch Studien, die zukünftig ins Feld übertragen werden sollen, durchgeführt. Aufgrund der unterschiedlichen Systemanforderungen von Labor- und Feldstudien wurden für die Durchführung der Studien verschiedene Messsysteme verwendet, welche in Abschnitt 4.2 näher erläutert werden. Für eine ausführliche Übersicht über die Vielzahl verfügbarer Geräte sei auf Curtin et al. (2007) und Ebner-Priemer & Kubiak (2007) verwiesen. Eine Darstellung der technologischen Entwicklungen der letzten Jahre, welche auf immer kleinere, belastungsfreie System abzielen, findet sich bei Patel et al. (2012) sowie Pantelopoulos & Bourbakis (2010).

Die wichtigsten Unterscheidungsmerkmale von Geräten zur Aufzeichnung physiologischer Daten sind nachfolgend dargestellt.

**Mobilität** Bei Geräten zur Erfassung physiologischer Daten kann zwischen mobilen und stationären Systemen unterschieden werden. Während bei stationären Systemen die Größe des Gerätes keine Rolle spielt, ist der Fokus bei der Entwicklung mobiler Systeme in erster Linie darauf gerichtet, dass diese im Alltag möglichst belastungsfrei getragen werden können und der jeweilige Träger in seinem Tagesablauf nicht behindert wird.

**Onlinefähigkeit** Ein wichtiges Unterscheidungsmerkmal bei der Aufzeichnung und Analyse physiologischer Daten stellt die Onlinefähigkeit der Geräte dar. Bei vielen Einsatzgebieten ist es nicht erforderlich, dass die aufgezeichneten Daten in Echtzeit analysiert werden können. In diesem Fall ist es lediglich notwendig, dass der verfügbare Speicher für die aufzuzeichnende Datenmenge ausreichend bemessen ist. Jedoch gibt es auch Anwendungsszenarien, in denen Daten in Echtzeit analysiert werden sollen. Dies erfordert, dass die Daten entweder direkt auf dem Gerät weiterverarbeitet werden oder aber über eine Schnittstelle an ein anderes Gerät wie beispielsweise einen Computer oder ein Handy übertragen werden.

**Verfügbarkeit** Viele für Forschungszwecke entwickelte Geräte sind nicht kommerziell verfügbar. Entweder handelt es sich hierbei um Geräte, die nur für den Eigenbedarf entwickelt wurden oder aber um Prototypen, welche noch nicht ausreichend entwickelt sind, um sie dem Endkunden anzubieten. Andere Hersteller wiederum bieten genau auf den Kunden zugeschnittene Systeme an, die jedoch zumeist auch nur in geringer Stückzahl erhältlich sind. Demgegenüber stehen für den Massenmarkt entwickelte Endprodukte, die auch in größerer Anzahl erworben werden können.

**Medizinproduktzulassung** Je nachdem, in welchem Gebiet die Messgeräte eingesetzt werden sollen, sind verschiedene Vorschriften zu beachten. So benötigen Geräte, welche in der medizinischen Diagnostik zum Einsatz kommen sollen, eine Zulassung als Medizinprodukt gemäß der Norm DIN EN ISO 13485. Bei Geräten, die lediglich zu nichtmedizinischen Forschungszwecken eingesetzt werden, ist eine derartige Zulassung nicht erforderlich.

**Offenheit** Viele kommerzielle Systeme erlauben keinen oder nur eingeschränkten Zugriff auf die Rohdaten, die während einer Messung aufgezeichnet werden. Dies liegt zumeist darin begründet, dass zur Analyse der Daten eine ebenfalls kommerziell erhältliche Software genutzt werden soll. Des Weiteren sind häufig bereits in der Hardware Komponenten verbaut, die eine Filterung des Signals vornehmen. Auch gibt es Geräte, die bereits bei der Aufnahme eine Feature-Extraktion vornehmen und nur diese extrahierten Features (wie beispielsweise aus einem EKG-Signal berechnete Herzrate) speichern bzw. übertragen.

**Erweiterbarkeit** Während es Geräte gibt, welche lediglich auf die Aufzeichnung eines einzelnen Signals ausgelegt sind, gibt es auf der anderen Seite flexible Systeme, welche individuell an die Studienanforderungen angepasst werden können. Diese ermöglichen die synchrone Aufzeichnung einer Vielzahl physiologischer Signale mit nur einem System.



**Abbildung 3.1:** Unterschiedliche Arten mobiler Polygraphen zur Aufzeichnung physiologischer Signale. Links: Varioport-e ohne Medizinproduktzulassung, Rechts: Somnoscreen mit Medizinproduktzulassung

In Abbildung 3.1 sind beispielhaft unterschiedliche Arten von Sensoren zur Aufzeichnung physiologischer Signale dargestellt.

## 3.2 Messung von Emotionen

Emotionen können sich auf unterschiedliche Weise äußern. Aus der Art ihres Auftretens können verschiedene Komponenten abgeleitet werden. Nach Schmidt-Atzert (1981); Izard (1994) setzen sich Emotionen aus

folgenden drei Komponenten zusammen:

1. Die *subjektiv wahrgenommene Emotion*, die den emotionalen Zustand betrifft, den eine Person selbst wahrnimmt bzw. beschreibt.
2. Das *emotionale Verhalten*, das im Zuge einer Emotion gezeigt wird.
3. Die *emotionalen physiologischen Reaktionen* im Gehirn und dem Nervensystem, die auf emotionale Stimuli zurückgeführt werden können

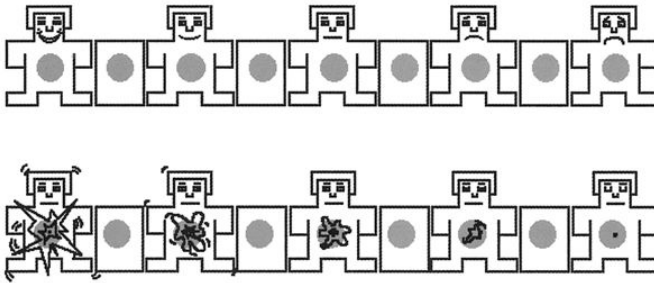
Zur Erfassung dieser einzelnen Komponenten bieten sich verschiedene Methoden an. Während die subjektiv wahrgenommenen Emotionen oftmals über Fragebögen erfasst werden, kann die Messung der psychophysiologischen Reaktionen über Biosignale erfolgen. Zur Analyse des emotionalen Verhaltens wiederum werden üblicherweise externe Beobachter eingesetzt.

#### 3.2.1 Erfassung des emotionalen Erlebens

Bei der Messung des emotionalen Erlebens geht es um die Erfassung des subjektiv wahrgenommenen emotionalen Zustands einer Person. Um diesen für Außenstehende zugänglich zu machen, werden in der Regel Fragebögen oder Selbstberichte genutzt. In der Praxis hat es sich bewährt, 5- oder 7-stufige Skalen zu verwenden (z. B. Lozano et al., 2008).

Im Rahmen dieser Arbeit wird zur Erfassung des wahrgenommenen emotionalen Zustands das Self-Assessment-Manikin (SAM) verwendet. Das SAM nach Lang (1980) ist ein bildbasierter Fragebogen, mithilfe dessen der emotionale Zustand einer Person auf einer bis zu 9-stufigen Skala hinsichtlich der Kriterien Valenz, Erregung sowie Dominanz erfasst werden kann. In einer von Bradley & Lang (1994) durchgeführten Studie konnte gezeigt werden, dass dieser Fragebogen sehr gut zur schnellen Evaluation von Valenz und emotionaler Erregung geeignet ist. Abbildung 3.2 zeigt die Skalen für Valenz und Erregung.

Neben diesem Fragebogen gibt es eine Reihe weiterer standardisierter Verfahren zur Erfassung des emotionalen Zustands wie beispielsweise die Positive Affect Negative Affect Scale (PANAS, Watson et al., 1988), bei welcher 20 emotionale Wörter auf einer 5-stufigen bipolaren Ratingskala bewertet werden müssen. Der Vorteil des SAM-Fragebogens besteht darin,



**Abbildung 3.2:** SAM-Fragebogen für Valenz (oben) und Erregung (unten) (Lang et al., 2005)

dass mit nur zwei Fragen in sehr kurzer Zeit ein für diese Arbeit hinreichend genaues Bild des Zustands einer Person geben kann. Zudem ist aufgrund der Sprachfreiheit des Fragebogens auszuschließen, dass einzelne Items von verschiedenen Personen unterschiedlich interpretiert werden.

### 3.2.2 Analyse des beobachtbaren Verhaltens

Emotionen manifestieren sich im Ausdruck einer Person auf vielfältige Weise. So kann sich aufgrund einer Emotion beispielsweise der Gesichtsausdruck oder die Körperhaltung oder aber auch der Tonfall beim Sprechen verändern. Die Beurteilung des emotionalen Verhaltens einer Person erfolgt üblicherweise unter Zuhilfenahme von Audio- und / oder Videoaufzeichnungen. Untersucht werden hierbei in erster Linie die Mimik sowie teilweise auch der Stimmausdruck der beobachteten Personen.

Bei der Mimik-Analyse dient in der Regel das Facial Action Coding System (Ekman & Friesen, 1977) als Grundlage, nach welchem das Gesicht in 44 sogenannte Aktionseinheiten (*action units*) unterteilt werden kann. Die zuverlässigste Methode zur Analyse des emotionalen Ausdrucks stellt die manuelle Annotation durch Experten dar. Jedoch ist diese mit einem sehr hohem Aufwand verbunden – die Annotation eines einminütigen Videos durch Experten dauert ungefähr eine Stunde (Donato et al., 1999). Die computerbasierte Mimik-Analyse hingegen ist zum aktuellen Zeitpunkt noch sehr anfällig für Störungen und es gibt nur wenige Systeme, die auch



unter Echtzeit-Bedingungen funktionieren.

Zwar sind die Erkennungsraten bei der Analyse von Audiodaten etwas geringer als bei der Mimik-Analyse, jedoch ist diese unabhängiger von Umgebungseinflüssen und zudem auch besser für Echtzeit-Anwendungen geeignet. Ein ausführlicher Überblick über audio- und videobasierte Systeme zur Emotionserkennung findet sich bei Zeng et al. (2009).

#### 3.2.3 Psychophysiologische Korrelate von Emotionen

Bei der Untersuchung psychophysiologischer Korrelate von Emotionen geht es darum herauszufinden, inwiefern sich physiologische Muster für unterschiedliche emotionale Zustände unterscheiden.

Eine der zuverlässigsten Methoden zur Messung des emotionalen Zustandes einer Person stellt die Messung des Stresshormons Cortisol dar (Grillon et al., 2006). Die Messung erfolgt üblicherweise über Speicheltests. Da eine Auswertung der Speichelproben in Echtzeit derzeit noch nicht möglich ist, ist diese Methode für Anwendungen, bei denen eine geringe Latenz bei dem Feedbackprozess erforderlich ist, nicht geeignet.

Statt der direkten Messung der biochemischen Prozesse des Körpers können auch die Auswirkungen des sympathischen und parasympathischen Nervensystems auf den Körper gemessen werden. So kann ein emotionaler Stimulus beispielsweise emotionales Schwitzen, Veränderungen der Herzrate oder der Atmung mit sich bringen. Im Gegensatz zum Hormonspiegel können diese physiologischen Signale in Echtzeit mittels entsprechender Sensorik ausgewertet werden.

Im Folgenden wird dargestellt, wie sich die im Rahmen dieser Arbeit verwendeten physiologischen Signale in Abhängigkeit von der Aktivierung des sympathischen und parasympathischen Nervensystems verhalten.

##### 3.2.3.1 Kardiovaskuläre Aktivität

Emotionale Erregung spiegelt sich in verstärkter sympathischer bzw. verringerter parasympathischer Aktivität wieder. Die Herzrate reflektiert sowohl sympathische als auch parasympathische Aktivitäten. Während sympathische Aktivierung zu einem Anstieg der Herzrate führt, spiegelt sich parasympathische Aktivität in einer Verlangsamung der Herzrate wider

(Berntson et al., 2007). Da emotionale Erregung in der Regel mit einer sympathischen Aktivierung einhergeht, führt dies auch zu einer steigenden Herzrate. Zusätzlich kann die HRV Aufschlüsse über das Zusammenspiel von sympathischer und parasympathischer Aktivität geben. Der Grad dieses Zusammenspiels kann als Indikator für die Emotionsregulationsfähigkeiten einer Person gesehen werden (Appelhans & Luecken, 2006). Sympathische Aktivität spiegelt sich vor allem in niedrigeren LF-Frequenzbändern wider, wohingegen hochfrequente Änderungen im HF-Frequenzbereich als Maß für hohe parasympathische Aktivität gesehen werden (Berntson et al., 2007) können. Das Zusammenspiel von sympathischem und parasympathischem Nervensystem wird durch das LF/HF-Ratio ausgedrückt.

### 3.2.3.2 Elektrodermale Aktivität

Im Gegensatz zum Herz werden die Schweißdrüsen in erster Linie vom sympathischen Nervensystem innerviert. Zwar ist bei der Regulierung der EDA mit Acetylcholin ein Neurotransmitter beteiligt, der üblicherweise der parasympathischen Aktivität zugeordnet wird, jedoch spiegelt sich das nicht in den physiologischen Reaktionen wider (Dawson et al., 2007). Unter normalen Umgebungsbedingungen konnte von Wallin (1981) eine hohe Korrelation zwischen sympathischer Aktivierung und den Hautleitwertreaktionen festgestellt werden.

Studien über phasische elektrodermale Reaktionen werden meistens nach dem Stimulus-Response-Schema durchgeführt, d. h. es werden Hautleitwertreaktionen als Response auf einen zuvor gezeigten Stimulus untersucht. Zahlreiche Studien haben gezeigt, dass die Amplitude der SCRs mit zunehmender emotionaler Erregung steigt und somit einen guten Indikator für emotionale Erregung darstellt (z. B. Lang et al., 1993; Cuthbert et al., 1996). Voraussetzung für derartige Analysen ist jedoch das Vorhandensein eines konkreten Stimulus. Stimuli die einem hohen emotionalen Erregungsniveau (egal ob positiv oder negativ) zugeordnet wurden, führten bei Studien mit akustischen Stimuli aus emotionalen Geräuschen (Bradley & Lang, 2000) und Musik (Gomez & Danuser, 2004) sowie bei Videostimuli (Christie & Friedman, 2004) zu größeren elektrodermalen Reaktionen als neutrale Stimuli.

Auch bei tonischen elektrodermalen Reaktionen konnten Zusammenhänge zu emotionalen Zuständen nachgewiesen werden. So zeigte sich

beispielsweise bei Gross & Levenson (1993); Gross (1998) eine Erhöhung des SCL während des Betrachtens emotionaler Filme.

#### 3.2.3.3 Muskelaktivität

Für die Messung der Muskelaktivität stehen zwei unterschiedliche Methoden zur Verfügung: Zum einen kann mittels Elektromyographie (EMG) die direkte Aktivität eines Muskels erfasst werden, zum anderen besteht die Möglichkeit, die Auswirkungen der Muskelaktivität in Form der ausgeübten Kraft zu erfassen.

Wie bereits in Abschnitt 3.2.2 dargestellt, können aus dem Gesichtsausdruck einer Person Informationen über deren emotionalen Zustand abgeleitet werden. Alternativ zur Videoanalyse kann dies auch über die Messung der Gesichtsmuskelaktivität (facial EMG) erfolgen (z. B. Dimberg, 1990). Jedoch erfordert dies die Platzierung zahlreicher Elektroden im Gesicht der zu beobachtenden Person.

Wie Arbeiten von Reynolds & Reynolds (2001) und Dennerlein et al. (2003) zeigen, spiegeln sich emotionale Zustände ebenfalls in der Muskelkraft wider, welche bei der Bedienung einer Computermaus angewendet wird: Höhere Frustration führt dazu, dass mehr Kraft auf eine Maus ausgeübt wird.

## 3.3 Affective Computing

In der Interaktion zwischen Menschen spielen Emotionen eine wichtige Rolle. Die Fähigkeit, den emotionalen Zustand der uns umgebenden Personen zu erkennen und entsprechend darauf zu reagieren ist ein wesentlicher Bestandteil zwischenmenschlicher Kommunikation. Derartige Informationen auch in der Mensch-Maschine-Interaktion zur Verfügung zu stellen ist Kernanliegen des Bereichs des *Affective Computings*. Der Begriff geht zurück auf Rosalind Picard (1995), die diesen wie folgt definierte:

“ computing that relates to, arises from, or deliberately influences emotion ”

Dieser Abschnitt gibt zunächst eine Übersicht über die Anwendungsgebiete des Affective Computing. Darauf folgend wird ein Überblick über Systeme

zur Klassifikation von Emotionen gegeben. Zuletzt werden Softwarelösungen zur Verarbeitung multimodaler Sensordaten vorgestellt, welche zum Rapid Prototyping von affektiven Anwendungen genutzt werden können.

### 3.3.1 Anwendungsfelder

Für affektsensitive Systeme gibt es zahlreiche Anwendungsfälle. Neben den affektiven Lernsystemen, auf welche in Abschnitt 3.4.1 detaillierter eingegangen wird, gibt es eine Vielzahl weiterer Möglichkeiten, bei welchen Informationen über den emotionalen Zustand eines Nutzers genutzt werden, um auf diesen genauer einzugehen. So wurden bereits zahlreiche Spiele entwickelt, welche auf den emotionalen Zustand eines Spielenden reagieren (Gilleade et al., 2005; Liu et al., 2009; Sykes & Brown, 2003; Jerčić et al., 2012). Im Bereich der Autismusforschung können affektive Systeme helfen, um Autisten dabei zu unterstützen ein besseres Verständnis für ihre eigenen und für fremde Emotionen zu entwickeln (R. Picard, 2009; Gay et al., 2013). Auch als Unterstützung zur Selbstreflektion können affektive Systeme als Hilfe dienen (Hernandez et al., 2013). Auf dem Gebiet der Robotik haben affektive Systeme ebenfalls Einzug gehalten (Breazeal, 2003; Fong et al., 2003; Rani et al., 2006; Schaaff & Schultz, 2009), um die Interaktion zwischen Mensch und Roboter menschenähnlicher zu gestalten. Selbst bei Fahrerassistenzsystemen findet der emotionale Zustand einer Person inzwischen Beachtung (Eyben et al., 2010; Schneider et al., 2012), da das frühzeitige Erkennen von Müdigkeits- und Stresszuständen erheblich zur Sicherheit des Fahrers beitragen kann.

### 3.3.2 Klassifikation von Emotionen

Der Entwurf affektiver Systeme erfordert Methoden zur Erkennung emotionaler Informationen, um diese in die Interaktion mit dem Computer zu integrieren. Zahlreiche Studien beschäftigen sich daher mit der automatischen Klassifikation von Emotionen. Bei nahezu allen Studien wird nach dem in Abbildung 3.3 dargestellten Schema vorgegangen.

Zur Induktion emotionaler Zustände gibt es verschiedene Ansätze. Insbesondere bei der Erkennung von Emotionen anhand des beobachtbaren Verhaltens kommen oft gespielte Emotionen zum Einsatz, bei welchen



**Abbildung 3.3:** Schritte bei der Emotionsklassifikation

die Versuchspersonen sich in einen bestimmten emotionalen Zustand hineinversetzen sollen (z. B. Castellano et al., 2008). Da diese jedoch nicht zwangsläufig natürlichen Emotionen entsprechen, werden oftmals emotionale Stimuli wie Filme (Ivonin et al., 2013), Musik (Eich et al., 2007) oder Spiele (Schaaff, Degen et al., 2012) verwendet, um unterschiedliche emotionale Zustände bei den Versuchspersonen auszulösen. Zudem gibt es mit dem *International Affective Picture System* (IAPS, Lang et al., 2005) standardisierte Bilder, die speziell zum Zweck der Emotionsinduktion zusammengestellt wurden. Die meisten der durchgeführten Studien werden unter kontrollierten Bedingungen im Labor durchgeführt, da dies eine sehr gute Kontrolle der Umweltbedingungen erlaubt. Unerwartete Ereignisse können weitestgehend ausgeschlossen werden und haben dadurch keinen Einfluss auf die Messung. Studien im Feld hingegen ermöglichen zwar eine Beobachtung der Versuchspersonen unter realen Bedingungen, jedoch stellt sich die Kontrolle der Umgebungsvariablen, welche die Messung beeinflussen, oft als sehr schwierig dar.

Während der Durchführung psychophysiologischer Studien treten oftmals Artefakte auf – sei es aufgrund von Bewegung, technischen Störungen oder anderen Umgebungseinflüssen. Im Rahmen der Vorverarbeitung werden diese Artefakte mittels unterschiedlicher Filter reduziert. Des Weiteren kann die Vorverarbeitung eine Normalisierung der Daten beinhalten, welche dazu dient, Daten unterschiedlicher Versuchspersonen vergleichbar zu machen. So wird beispielsweise bei Mandryk & Atkins (2007) eine Normalisierung der physiologischen Signale auf einen Wertebereich von 0 bis 100 vorgenommen.

Im nächsten Schritt müssen aus den vorverarbeiteten Signalen Features extrahiert werden, welche mit dem emotionalen Zustand einer Person in Verbindung stehen. Die Möglichkeiten bei der Feature-Extraktion sind vielfältig und reichen von einfachen statistischen Features wie dem Signalmittelwert in einem bestimmten Fenster bis zu Frequenzspektren, wie

sie bei der HRV-Analyse gebräuchlich sind. Die Zahl der in den einzelnen Studien verwendeten Features variiert stark: während Benovoy et al. (2008) aus EKG, PPG, EDA, Temperatur und Atmung insgesamt 225 unterschiedliche Features berechnen, kommen andere Studien (z. B. Choi & Woo, 2005) mit weniger als fünf Features aus.

Bei der anschließenden Feature-Reduktion werden die Features selektiert, welche für die Klassifikation die größte Bedeutung haben, da nicht alle extrahierten Features zwangsläufig mit dem emotionalen Zustand korreliert sein müssen. Hierfür kommen verschiedene Algorithmen wie Sequential Floating Forward Selection (SFFS, Oliveira et al., 2011), die Lineare Diskriminanzanalyse (LDA, Zhou et al., 2011) oder die Principal Component Analysis (PCA, Valenza et al., 2012) zum Einsatz.

Den letzten Schritt stellt die eigentliche Klassifikation dar. Zum Einsatz kommen hier insbesondere der  $k$ -Nearest Neighbors Algorithmus, Diskriminanzanalyse, Neuronale Netze, Entscheidungsbäume und Support-Vector-Machines. Ein Überblick über Studien, bei welchen die unterschiedlichen Klassifikatoren zum Einsatz kommen, findet sich bei Novak et al. (2012). Je nach verwendetem Verfahren konnten Klassifikationsgenauigkeiten von über 90 % erreicht werden (z. B. Kim & André, 2008; Sakr et al., 2010). Ein Vergleich der Klassifikationsgenauigkeiten ist jedoch schwierig, da in verschiedenen Studien zwischen unterschiedlich vielen Klassen differenziert wird. Zudem findet die Klassifikation teilweise personenabhängig, teilweise aber auch personenunabhängig statt. Grundsätzlich liegt die Genauigkeit bei gespielten Emotionen höher als bei induzierten Emotionen (D'Mello & Kory, 2012).

Bei den meisten Studien kommen multimodale Klassifikatoren zum Einsatz, welche mehr als ein Biosignal berücksichtigen. Eine Meta-Analyse von D'Mello & Kory (2012) ergab jedoch, dass die Steigerungen, die durch die Multimodalität erzielt werden, mit einer durchschnittlichen Verbesserung von 8,12 % gegenüber einer unimodalen Klassifikation zwar deutlich erkennbar sind, jedoch niedriger ausfallen als erwartet. Insbesondere in Hinblick darauf, dass jeder zusätzliche Sensor zur Datenakquise eine zusätzliche Belastung für den Träger bedeutet, kann daher je nach Einsatzgebiet ein unimodales System zur Emotionsklassifikation sinnvoller sein.

Für eine ausführliche Übersicht über Studien, bei welchen Biosignale zur Emotionsklassifikation verwendet wurden, sei zusätzlich auf van den Broek et al. (2009, 2010, 2011) verwiesen.

#### 3.3.3 Software-Lösungen zur Verarbeitung multimodaler Daten

Es stehen bereits einige kommerzielle sowie nichtkommerzielle Open-Source-Lösungen zur Aufzeichnung und Analyse physiologischer Signale zur Verfügung. Bei den meisten kommerziellen Systemen handelt es sich um Closed-Source-Produkte, die an die Verwendung spezieller Sensoren gebunden und nicht flexibel erweiterbar sind. Zudem sind Messmethodik und die verwendeten Algorithmen oft nur unzureichend dokumentiert. Jedoch führt das zunehmende Interesse an Software-Lösungen zur Verarbeitung von Multisensordaten dazu, dass immer mehr Open-Source-Lösungen verfügbar werden. Die bekanntesten Frameworks werden im Folgenden kurz vorgestellt.

Bei der weitverbreiteten BCI2000-Software (Schalk et al., 2004) handelt es sich um eine sehr umfangreiche Software, die speziell für Brain-Computer-Interfaces entwickelt wurde. Die Software kann sowohl zur Präsentation von Stimuli als auch zur Aufzeichnung und Online-Auswertung von Daten verwendet werden. Der Schwerpunkt dieser Software liegt auf der Verarbeitung von EEG-Daten. Zwar ist eine Erweiterung um weitere Datenquellen möglich, jedoch beschränkt sich der Fokus der Anwendung derzeit auf die EEG-Datenverarbeitung.

Mit Frameworks wie ICARE (Bouchet et al., 2004), CrossWeaver (Sinha & Landay, 2003), ICON (Dragicevic & Fekete, 2001) oder OpenInterface (Lawson et al., 2010) stehen eine Reihe weiterer Systeme zum Rapid Prototyping multimodaler Systeme zur Verfügung. Jedoch sind nicht alle dieser Frameworks problemlos erweiterbar und es fehlen Komponenten zur Erfassung und Verarbeitung physiologischer Daten. Für weitere Informationen über Frameworks zur Verarbeitung multimodaler Daten sei auf Dumas et al. (2009) verwiesen.

## 3.4 Technologiestütztes Lernen

Der nachfolgende Abschnitt dient der Darstellung, in welcher Form Informationen über den emotionalen Zustand einer Person mittels physiologischem Feedback bereits in Lerntechnologien integriert worden sind. Bei den affektiven Lernsystemen handelt es sich in erster Linie um Lernumgebungen, die darauf abzielen, dem Nutzer ein bestimmtes Domänenwissen wie bei-

spielsweise aus der Physik oder der Medizin zu vermitteln. Diese werden im nachfolgenden Kapitel vorgestellt. Zum Training des Arbeitsgedächtnisses existieren derzeit noch keine Anwendungen, welche den emotionalen Zustand eines Nutzers berücksichtigen. Aus diesem Grund ist in Abschnitt 3.4.2 ein Überblick über verschiedene Anwendungen zum Training des Arbeitsgedächtnisses gegeben, bei welchen jedoch kein physiologisches Feedback verwendet wird. Zuletzt erfolgt ein Überblick über Anwendungen zum Training der Emotionsregulationsfähigkeiten.

### 3.4.1 Affektive Lernumgebungen

In den vergangenen Jahren haben sich affektive Lernumgebungen – sogenannte *Affective Tutoring Systems* – zum Forschungsgegenstand entwickelt. Im Gegensatz zu den *Intelligent Tutoring Systems* berücksichtigen diese auch Information über den emotionalen und / oder motivationalen Status eines Lernenden. Dies kann beispielsweise durch eine Anpassung des Lernstoffes an den emotionalen Zustand eines Nutzers oder aber auch durch eine natürlichere und motivierendere Interaktion mit dem Lernenden geschehen. Durch Anpassung der Lernsoftware an diese emotionalen Informationen soll eine Maximierung des Lernerfolgs erfolgen (D’Mello et al., 2007; Sarrafzadeh et al., 2008; Woolf et al., 2009). Nachfolgend sind die bekanntesten affektiven Lernumgebungen kurz dargestellt.

#### AutoTutor

Bei dem *AutoTutor* (D’Mello et al., 2008) handelt es sich um ein Physik-Lernsystem, welches sprachbasiert mit dem Lernenden interagiert und in dieser Interaktion auf dessen emotionalen Zustand reagiert. Ziel ist es, Motivation und Engagement bei dem Lernenden auf einem möglichst hohen Niveau zu halten. Das System erfasst die Emotionen Langeweile, Frustration und Verwirrung basierend auf dem Gesichtsausdruck und der Körperhaltung eines Nutzers. Bei den berücksichtigten Emotionen wurde von den Basisemotionen nach Ekman ausgegangen und solche Emotionen vernachlässigt, welche für den Lernenden nicht relevant sind. Je nach detektiertem emotionalen Zustand werden Sprache und Ausdruck eines Avatars, welcher mit dem Nutzer interagiert, an dessen Zustand angepasst.

Eine Evaluation des AutoTutors zeigte, dass eine Version des AutoTu-



tor mit emotionalem Support zu einem größeren Lernerfolg führte, als die gleiche Version des Systems ohne den emotionalen Support. Jedoch profitierten insbesondere Studenten mit geringerem Wissen von dem emotionalen Support, während bei Studenten mit einem hohen Wissensstand keine Steigerung des Lernerfolgs zu beobachten war. Systeme wie der AutoTutor können demnach insbesondere bei Studenten, welche einen geringen Wissensstand in der zu lernenden Domäne aufweisen, von Nutzen sein (D'Mello et al., 2010).

#### Easy with Eve

Ähnlich wie bei dem AutoTutor, dienen bei dem affektsensitiven Mathe-Lernsystem *Easy with Eve* (Alexander et al., 2006; Sarrafzadeh et al., 2008) die Basisemotionen nach Ekman als Grundlage für die Interaktion mit dem Lernenden. Diese werden mithilfe von Kamerasystemen aus Gestik und Mimik des Lernenden ermittelt und in sieben verschiedene Zustände klassifiziert. Die Reaktion auf die erkannten Nutzerzustände wird mittels eines Case-Based-Reasoning Systems ermittelt. Hierbei wird versucht, die Korrespondenzhistorie möglichst gut in die Kommunikation zu integrieren. Zur Nutzerinteraktion wird auch bei diesem System ein Avatar (Eve) genutzt.

#### Edu-Affe-Mikey

Im Gegensatz zu Easy with Eve und dem AutoTutor kommt das System *Edu-Affe-Mikey* (Alepis et al., 2006, 2007) ohne zusätzliche Sensorik aus. Das für Medizinstudenten entwickelte System verwendet als Eingabeparameter lediglich Features, welche aus dem Mikrofon und der Tastatur des Computers extrahiert wurden. Zur Identifikation von Ereignissen, welche eine Veränderung des emotionalen Zustands ausdrücken, wurde Expertenwissen in das System integriert. Wird eines dieser Ereignisse vom System erkannt, wird die dem Ereignis zugeordnete Reaktion des Systems ausgelöst. Wie auch beim AutoTutor und Easy with Eve erfolgt die Interaktion mit dem System über einen Avatar.

#### Weitere affektive Lernsysteme

Bei den oben vorgestellten Systemen handelt es sich um die bekanntesten affektiven Lernsysteme, welche sich unterschiedlicher Signale zur Erkennung

von Emotionen bedienen. Zudem gibt es eine Reihe weiterer dialogbasierter Systeme, wie beispielsweise das System *ITSPoke* (Litman & Silliman, 2004) – welches basierend auf den Nutzereingaben auf Unsicherheit des Lernenden reagieren soll – oder die videobasierten Systeme *EMASPEL* (Ben Ammar et al., 2010) und *AITs* (Nkambou, 2006). Das große Forschungsinteresse auf dem Gebiet der affektiven Lernsysteme zeigt, dass dieses Thema auch in Zukunft von Bedeutung sein wird.

### 3.4.2 Training des Arbeitsgedächtnisses

Wie in Abschnitt 2.2.3.2 eingeführt, gibt es eine Reihe standardisierter Methoden zum Training des Arbeitsgedächtnisses. Nach Klingberg (2010) liegt ein essenzieller Erfolgsfaktor des Trainings darin, dass die Aufgabe adaptiv gestaltet ist und sich während des Trainings individuell an die Leistung eines Nutzers anpasst. Des Weiteren sollte die gesamte Trainingsdauer bei mindestens 10 Stunden liegen. Diese Faktoren konnten auch im Rahmen anderer Studien bestätigt werden (Jaeggi et al., 2008; Holmes et al., 2009).

Die in Abschnitt 2.2.3.2 beschriebene  $n$ -Back-Task stellt eines der am weitesten verbreiteten Arbeitsgedächtnistrainings dar, welches in zahlreichen Studien zum Einsatz kam (z. B. Jaeggi et al., 2008; Holmes et al., 2009; Klingberg, 2010; Jaeggi et al., 2010; Harbison et al., 2011; Heinzl et al., 2013). Ein großer Vorteil der  $n$ -Back-Task besteht in der Vielzahl der zur Verfügung stehenden Adaptionsmechanismen:

- Eine *Variation des Parameters  $n$*  stellt den wohl bedeutendsten Adaptionsmechanismus dar. Dieser Parameter gibt an, mit welchem zurückliegenden Element das aktuelle Element verglichen werden muss. Je größer  $n$ , umso mehr Informationen muss eine Person im Gedächtnis behalten und umso höher ist folglich auch die Schwierigkeit.
- Die sogenannten *Lures* (Verlockungen) bezeichnen die Ähnlichkeit des aktuellen Elements zu benachbarten Elementen des um  $n$  zurückliegenden Elements. Bei einer 3-Back-Task, bei welcher zuvor die Buchstaben Y — Z — X präsentiert wurden, würde die Präsentation des Buchstaben Z als nächstes Element einem Lure von 1 entsprechen.

- Zuletzt können auch *zeitliche Parameter* die Schwierigkeit der Aufgabe bestimmen. Hier kann zum einen die Präsentationszeit eines Elementes variiert werden, aber auch die Zeit zwischen zwei angezeigten Elementen – das Interstimulus-Intervall (ISI).

In der Literatur finden sich verschiedene Adaptionstufen bei adaptiven  $n$ -Back-Tasks. Bei Jaeggi et al. (2003) erfolgt eine Anpassung des Parameters  $n$  im Bereich von 1 bis 4. Die Anpassung wird jeweils nach Ende einer Runde mit 20 Elementen vorgenommen. Eine Performance von über 95 % korrekten Antworten führt zu einer Erhöhung von  $n$  um 1 während bei einer Performance von unter 70 %  $n$  um 1 verringert wird.

Harbison et al. (2011) variieren während des Trainings sowohl die Lures als auch  $n$ . Während es für die Lures drei verschiedene Stufen gibt (Stufe 0: keine Lures, Stufe 1: Lures an Position  $n + 1$ , Stufe 2: Lures an Position  $n + 1$  und  $n - 1$ ), kann  $n$  zwischen 1 und 8 variieren. Eine Anpassung der Schwierigkeitsstufen nach oben wird ab einer Performance von 85 % vorgenommen, während eine Adaption nach unten ab 65 % erfolgt. Zunächst wird jeweils das Lure-Level angepasst. Ist hier der maximale Wert erreicht, erfolgt eine Anpassung von  $n$ . Die Anpassung der Performance erfolgt immer nach Ende einer Runde mit je 25 Elementen.

Heinzel et al. (2013) verwenden zur Adaption das ISI sowie  $n$ . Das ISI wird in Schritten von 200 ms zwischen 1000 ms und 1800 ms variiert, während  $n$  die Stufen 0 bis 5 annehmen kann. Bei der 0-Back-Bedingung besteht die Aufgabe darin, jeweils eine Taste zu betätigen, wenn das aktuelle Zeichen eine 0 zeigt. Die Adaption wird jeweils zwischen den Runden vorgenommen. Wird in der vorhergehenden Runde eine Performance von mindestens 80 % erreicht, wird die Schwierigkeit in der nächsten Runde um eins erhöht. Ähnlich wie auch bei Harbison et al. (2011) wird zunächst das ISI verringert und bei Erreichen der kürzesten Stufe anschließend  $n$  erhöht und das ISI wieder auf den längsten Wert zurückgesetzt.

Aufgrund des gestiegenen Interesses an Arbeitsgedächtnistrainings wurde in den letzten Jahren eine Vielzahl an kommerziellen und nichtkommerziellen Anwendungen zum Training des Arbeitsgedächtnisses entwickelt wie beispielsweise die Programme der Firma Cogmed (Cogmed, 2014), Jungle Memory (2014); NeuroNation (2014a). Auch bei diesen Trainingsprogrammen ist eine performancebasierte Adaption integriert, jedoch sind die Informationen über die genauen Adaptionmechanismen nicht offen

zugänglich.

Auch im mobilen Umfeld halten Apps zum Training des Arbeitsgedächtnisses Einzug. Im Google Play-Store finden sich zahlreiche kostenlose sowie kostenpflichtige Apps wie beispielsweise Neuro Nation (NeuroNation, 2014b), Cogmed (Pearson Education, Inc. Mobile, 2013), Memory Trainer (Urbian, 2014) oder Mind Games (Mindware Consulting, Inc, 2014), welche neben weiteren Gehirnfunktionalitäten auch das Arbeitsgedächtnis trainieren sollen.

### 3.4.3 Emotionsregulationstraining

Zum Training der Emotionsregulationsfähigkeiten mittels physiologischem Feedback kommen oftmals sogenannte Serious Games zum Einsatz. Hierunter sind Spiele zu verstehen, welche nicht nur dem Spielspaß dienen, sondern auch pädagogische Ziele verfolgen (Ritterfeld et al., 2009). Das Grundprinzip von Spielen, welche zum Training der Emotionsregulationsfähigkeit entwickelt wurden, besteht darin, dass das Spiel umso leichter wird, je entspannter ein Spieler ist. Erfolgreiche Regulation von Emotionen wird folglich belohnt.

Ein wichtiges Einsatzgebiet für derartige Trainings findet sich im klinischen Umfeld zur Behandlung emotionaler Verhaltensstörungen. So ist das Spiel RAGE-Control (Kahn et al., 2009) speziell auf die Behandlung jähzorniger Kinder ausgerichtet während der Fokus von Relax to Win (Sharry et al., 2003) auf der Behandlung von Angststörungen liegt. Ähnliche Spielprinzipien finden sich auch bei Entspannungstrainings wie Brainball (Hjelm, 2003) wieder.

Grundsätzlich ist Emotionsregulationstraining überall dort sinnvoll, wo Emotionen einen ungewollten Einfluss auf das Handeln nehmen können. Ein relativ neues Forschungsfeld stellen hierbei die Emotionsregulationsfähigkeiten von Händlern an Finanzmärkten dar. Speziell auf diese Zielgruppe angepasste Trainings können helfen, dass Emotionen besser gesteuert werden können und folglich bessere Handelsentscheidungen getroffen werden (Fenton-O’Creevy et al., 2012).



## 4 Systemkonzeption und -design

Die Integration von physiologischem Feedback in Lernanwendungen kann helfen, diese besser an den emotionalen Zustand eines Nutzers anzupassen. Auch kann die Reflexion über den emotionalen Zustand helfen, sich dieses Zustands bewusster zu werden und ihn gezielt zu beeinflussen. In diesem Kapitel wird daher zunächst in Abschnitt 4.1 die in dieser Arbeit betrachtete Problemstellung genauer erläutert und hieraus anschließend ein Ansatz entwickelt, mit welchem diese Problemstellung angegangen werden kann. Im Anschluss daran werden in den folgenden Abschnitten die einzelnen Komponenten des erarbeiteten Ansatzes vorgestellt.

Teile dieses Kapitels wurden bereits in abgeänderter Form in Schaaff, Degen et al. (2012); Schaaff, Müller et al. (2012) veröffentlicht. Die in Abschnitt 4.3 vorgestellte xAffect-Software steht zudem kostenlos unter <https://sourceforge.net/projects/xaffect> zum Download bereit.

### 4.1 Problemstellung und Ansatz

Der Erfolg beim Lernen wird von einer Vielzahl von Faktoren beeinflusst. Im Rahmen dieser Arbeit wird das Augenmerk insbesondere auf den Einfluss emotionaler Erregung gerichtet. Der Zusammenhang zwischen emotionaler Erregung und dem Lernerfolg findet sich sowohl im Yerkes-Dodson-Gesetz als auch im Flow-Modell wieder. Beide Modelle gehen davon aus, dass es einen optimalen emotionalen Erregungszustand für das Lernen gibt (siehe hierzu auch Abschnitt 2.2.2). Zur Erreichung dieses optimalen emotionalen Erregungszustandes werden in dieser Arbeit zwei unterschiedliche Ansätze verfolgt: die Regulation des emotionalen Erregungszustandes eines Nutzers an die Lernanwendung sowie die Adaption einer Lernanwendung an den Nutzerzustand.

Die Fähigkeit, die eigenen Emotionen gemäß den Anforderungen einer Lernumgebung zu regulieren, kann entscheidend dazu beitragen, dass

Lerninhalte besser verarbeitet und aufgenommen werden. Die Integration von Informationen über den emotionalen Zustand einer Person in dieses Training kann ihr an dieser Stelle helfen, im Rahmen eines Emotionsregulationstrainings ein besseres Bewusstsein für ihren emotionalen Zustand zu erhalten und zu lernen, die Emotionen gezielter zu regulieren.

Des Weiteren kann eine Adaption auf Seite der Lernanwendung erfolgen. Hierfür wird im Rahmen dieser Arbeit ein Arbeitsgedächtnistraining verwendet. Beim Arbeitsgedächtnistraining kann die Information über das emotionale Erregungsniveau genutzt werden, um während des Trainings die Schwierigkeit auf einem Level zu halten, welches dem optimalen emotionalen Erregungsniveau eines Nutzers entspricht. Hierdurch kann Frustration durch ein zu schwieriges Training bzw. Langeweile durch ein zu einfaches Training vermieden werden. Dies wiederum führt bei dem Nutzer zu einer höheren Motivation, das Training auch über einen längeren Zeitraum zu absolvieren, da Arbeitsgedächtnistraining erst nach mehreren Wochen Effekte auf die kognitiven Fähigkeiten hat (siehe hierzu auch Abschnitt 2.2.3).

In Abschnitt 3.2 wurden verschiedene Möglichkeiten dargestellt, Informationen über den emotionalen Zustand eines Nutzers zu erhalten. Die wohl aufwändigste und somit am wenigsten für echtzeitnahes Feedback geeignete Methode stellt die Beobachtung des Verhaltens durch eine andere Person dar. Computerbasierte Verfahren können hierbei zwar unterstützen, jedoch sind diese derzeit noch stark abhängig von externen Bedingungen wie einer optimalen Beleuchtungssituation oder der optimalen Ausrichtung des Gesichts auf eine Kamera.

Auch die Erfassung des emotionalen Erlebens mittels Fragebögen bringt Schwierigkeiten mit sich, da die hiermit verbundene Introspektion oft fehleranfällig ist und vielmehr Hypothesen über den eigenen Gefühlszustand reflektiert als den Gefühlszustand selbst (Nisbett & Wilson, 1977). Der Erfolg von Selbsteinschätzungen ist daher stark davon abhängig, dass die jeweiligen Personen sich ihres emotionalen Zustands bewusst sind und zudem in der Lage sind, diesen in dem vorgegebenen Erhebungsrahmen auch akkurat wiederzugeben. Aufgrund der subjektiven Komponente der Selbsteinschätzungen sind diese daher sehr anfällig für zufällige und systematische Messfehler (Coan & Allen, 2007). Des Weiteren spiegeln Selbsteinschätzungen immer nur den Moment rund um den Zeitpunkt der Reflexion wider. Ein kontinuierliches Erfassen der Selbsteinschätzung

ist nicht möglich. Liegt die Präsentation eines Fragebogens zu weit vom eigentlichen Zeitpunkt eines Ereignisses entfernt, können die Eindrücke in einer spezifischen Situation nicht mehr unverzerrt ins Gedächtnis zurückgerufen werden und unterliegen somit systematischen Fehlern (Robinson & Clore, 2002).

Die Analyse der physiologischen Reaktionen des Körper mittels psychophysiologischer Sensorik stellt das objektivste Verfahren dar, da es weitestgehend unanfällig gegenüber willentlicher Manipulation ist. Des Weiteren ermöglichen physiologische Messungen ein kontinuierliches Erfassen des Nutzerzustands und sind so für echtzeitnahes Online-Feedback über den emotionalen Zustand einer Person geeignet. Durch die zunehmende Miniaturisierung im Bereich der Sensortechnik ist die Messung physiologischer Signale inzwischen nahezu belastungsfrei mittels nicht-invasiver Sensorik möglich.

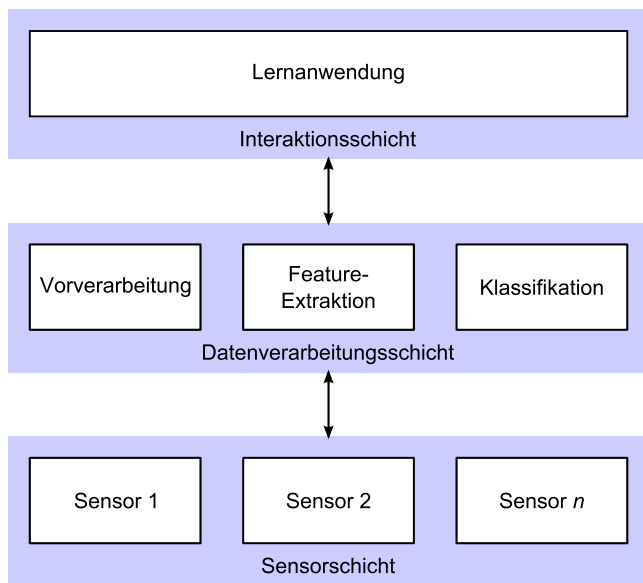
Zwar gibt es bereits Lernumgebungen, welche physiologisches Feedback integrieren, jedoch zielen diese auf das direkte Training des zu vermittelnden Stoffes ab. Zudem sind die verwendeten Sensorlösungen in den meisten Fällen nicht unter Alltagsbedingungen einsetzbar. Soll das Training regelmäßig erfolgen, ist die Einsatzfähigkeit unter Alltagsbedingungen jedoch ein wesentlicher Erfolgsfaktor für den Lernerfolg.

Ziel dieser Arbeit ist die Entwicklung eines Systems, welches die drei Themenbereiche Lernen, Emotionen und Sensortechnologien vereint. Psychophysiologische Sensortechnologien sollen genutzt werden, um den emotionalen Zustand einer Person mittels entsprechender Algorithmen messbar zu machen. Diese Informationen sollen durch die Integration in Lernanwendungen dazu dienen, den Lernerfolg zu steigern und die Motivation am Lernen zu erhöhen. Eine wichtige Anforderung an das System ist, dass das Lernen in den alltäglichen Umgebungen der Personen erfolgen kann, was besondere Anforderungen an die verwendeten Technologien stellt.

Abbildung 4.1 zeigt das Gesamtkonzept des im Rahmen dieser Arbeit vorgestellten Systems. Im Fokus des Konzepts steht, dass das System möglichst flexibel an sich verändernde Rahmenbedingungen hinsichtlich der zur Verfügung stehenden Sensoren und Algorithmen angepasst werden kann. Zudem soll es möglich sein, dass System in verschiedene Lernanwendungen zu integrieren.

Aus diesem Grund ist System aus drei Schichten aufgebaut: Die Sensor-schicht dient der Erfassung von Daten, die Rückschlüsse auf das emotionale





**Abbildung 4.1:** Aufbau des Gesamtsystems

Erregungsniveau eines Nutzers zulassen. Die Datenverarbeitungsschicht fungiert als Zwischenschicht zwischen der Lern- und Feedbackebene und den Sensoren. Ihre Aufgabe ist die Extraktion von Informationen aus den physiologischen Signalen sowie die hiermit verbundene Identifikation emotionaler Erregungszustände. Die Interaktionsschicht wiederum dient der Interaktion mit dem Benutzer und beinhaltet die eigentliche Lernanwendung.

Nachfolgend werden die drei Schichten sowie die damit verbundenen Anforderungen beschrieben.

### 4.1.1 Anforderungen an die Sensorschicht

Das zu entwickelnde System soll unter Alltagsbedingungen zum Einsatz kommen. Aus diesem Grund ist es erforderlich, dass die verwendeten Sen-

soren über einen längeren Zeitraum belastungsfrei getragen werden können. Stören diese den Nutzer während des Tragens kann dies dessen emotionalen Zustand ungewollt beeinflussen. Zudem wird ein Nutzer nicht bereit sein, ein störendes System über einen längeren Zeitraum zu verwenden. Belastungsfreiheit setzt ein geringes Gewicht der Sensoren sowie einen hohen Tragekomfort voraus. Auch sollten die verwendeten Sensoren beim Tragen nicht optisch auffallen. Das Anlegen und Abnehmen der Sensoren muss von dem Nutzer selbst vorgenommen werden können. Alternativ kann auch eine unsichtbare Integration der Sensoren in Alltagsgegenstände erfolgen, die während der Messung der Signale verwendet werden. Für eine möglichst akkurate Datenanalyse ist es erforderlich, dass die Sensoren in der Lage sind, physiologische Daten mit einer möglichst hohen Datenqualität zu erfassen und in Echtzeit an die Datenverarbeitungsschicht zu übertragen.

### 4.1.2 Anforderungen an die Datenverarbeitungsschicht

Im Bereich der mobilen Sensorik sind derzeit starke Veränderungen zu beobachten. Neue Systeme werden in immer kürzeren Abständen verfügbar. Auch bei den Algorithmen zur Datenanalyse gibt es häufig neue Entwicklungen. Aus diesem Grund sind Modularität und Konfigurierbarkeit zwei wichtige zu berücksichtigende Anforderungen an die Datenverarbeitungsschicht.

Hinsichtlich dieser beiden Anforderungen soll die Datenverarbeitungsschicht folgende Aktionen unterstützen:

- Neue Datenquellen wie Sensoren sollen mit wenig Aufwand an das System angebunden werden können.
- Die Integration neuer Algorithmen soll ohne Änderungen des bestehenden Systems möglich sein.
- Mehrere Datenquellen sollen zur gleichen Zeit an das System angebunden sein können.
- Die Datenverarbeitungsschicht soll unabhängig von der Lernsoftware funktionieren.

Zudem muss sichergestellt sein, dass das zu entwickelnde System in der Lage ist, emotionale Erregung aus den aufgezeichneten Signalen zu erkennen. Neben validierten Algorithmen zur Erkennung emotionaler Erregung

beinhaltet dies, dass das System sowohl die Rohdaten als auch die daraus abgeleiteten Features speichern kann, damit diese zur ex-post Analyse sowie zur weiteren Verbesserung der Algorithmen verwendet werden können.

Die Klassifikation der emotionalen Erregung soll ohne lange Wartezeiten für den Nutzer erfolgen. Nur so kann eine echtzeitnahe Einbindung der extrahierten Features in die Lernanwendung gewährleistet werden. Hierzu müssen die entwickelten Algorithmen in der Lage sein, emotionale Erregung mit einer möglichst geringen Latenz aus den aufgezeichneten Signalen abzuleiten. Jedoch ist es wichtig, dass bei der Merkmalsextraktion Reaktionszeiten der jeweiligen Signale berücksichtigt werden. Die Fensterlänge für die Merkmalsextraktion ist so zu wählen, dass das Signal ausreichend Informationen für die Analyse enthält. Die Algorithmen zur Ableitung der Features müssen in der Lage sein, die Daten ohne Verzögerung zu verarbeiten.

Zusätzlich muss die Datenverarbeitungsschicht geeignete Schnittstellen zur Kommunikation mit Lernanwendungen aus der Interaktionsschicht zur Verfügung stellen.

### 4.1.3 Anforderungen an die Interaktionsschicht

Neben generellen Anforderungen an Benutzerfreundlichkeit und korrekte Vermittlung der Lerninhalte, muss die Interaktionsschicht in der Lage sein, Informationen über den emotionalen Erregungszustand einer Person zu empfangen und hierauf entsprechend des Trainingsziels zu reagieren. Veränderungen des emotionalen Erregungszustands können entweder eine Verringerung oder Erhöhung des Schwierigkeitsniveaus zur Folge haben. Zudem muss die Lernanwendung in der Lage sein, Konfigurationsnachrichten mit der Datenverarbeitungsschicht auszutauschen.

Sofern im Rahmen der Lernanwendung Feedback über den Trainingszustand gegeben wird, soll zudem der emotionale Erregungszustand einer Person bei der Feedbackpräsentation Berücksichtigung finden.

## 4.2 Auswahl geeigneter Biosignale und Messsysteme

Wie in Abschnitt 3.3.2 deutlich wurde, gibt es eine Vielzahl von Signalen aus denen Informationen über den emotionalen Zustand einer Person

abgeleitet werden können. Jedoch genügen nicht alle den in Abschnitt 4.1.1 definierten Anforderungen.

Messungen des Stresshormons Cortisol können zwar ein sehr gutes Bild über den emotionalen Zustand einer Person geben, jedoch sind diese aktuell noch mit langen Auswertezeiten verbunden, so dass diese für echtzeitnahes Feedback ungeeignet sind. Zudem ist die Latenz bei den Veränderungen des Hormonspiegels von mehreren Minuten nicht für die Online-Datenanalyse geeignet. Bei der Analyse von Sprache besteht das Problem darin, dass insbesondere in Lernumgebungen nicht kontinuierlich Sprachsignale ausgesendet werden. Zur Analyse der Mimik anhand von Videodaten sind eine sehr hohe Bildqualität und optimale Beleuchtungsbedingungen erforderlich, die unter Alltagsbedingungen nicht gewährleistet werden können. Zudem muss sich das Gesicht immer im Bereich der Kamera befinden. Von daher kann diese Analyse nur für Anwendungen durchgeführt werden, bei welchen ein Nutzer ständig in die Richtung der Kamera schaut. Aus den genannten Gründen werden die genannten Signale in dieser Arbeit nicht weiter untersucht.

Auch bei den physiologischen Daten ist nicht jedes Signal für die unauffällige und belastungsfreie Analyse geeignet. So ist es bei der Messung von fEMG und EEG erforderlich, Sensoren auf dem Kopf bzw. im Gesicht zu platzieren, was ebenfalls nicht mit den vorher definierten Anforderungen vereinbar ist. Für die vorliegende Arbeit wird der Fokus daher auf kardiovaskuläre und eklektrodermale Signale gelegt, da für diese Signale Sensoren verfügbar sind, die unauffällig und belastungsfrei über einen längeren Zeitraum getragen werden können. Zusätzlich wird im Rahmen dieser Arbeit die Muskelkraft untersucht, die auf die Maus ausgeübt wird. Der große Vorteil dieses Signals liegt darin, dass der Nutzer für die Messung keine zusätzlichen Sensoren am Körper tragen muss. Jedoch kann dieses Signal nur in Szenarien genutzt werden, in denen der Nutzer einen Computer mit einer Maus bedient.

Im Folgenden werden die einzelnen Systeme dargestellt, die im Rahmen der in dieser Arbeit durchgeführten Studien zur Erfassung und Aufnahme der Sensordaten verwendet wurden. Neben den verwendeten kommerziellen Systemen wurde zur Erfassung der Kraft, welche auf die Maustaste ausgeübt wird, ein eigenes System entwickelt, da für diesen Zweck keine geeigneten Systeme erhältlich sind.

### 4.2.1 Kommerzielle Systeme zur Messung physiologischer Signale

Zur Erfassung der physiologischen Signale im Rahmen der durchgeführten Studien kamen verschiedene kommerziell verfügbare Geräte zur Aufzeichnung der Daten zum Einsatz. Diese werden im Folgenden beschrieben.

#### Labormesssystem

Das Labormesssystem wurde in Zusammenarbeit der Forschungsgruppe hiper.campus und dem ITIV am KIT entwickelt. Schwerpunkt der Entwicklung war die Erstellung eines modularen Systems zur Aufnahme physiologischer Parameter in Laborstudien. Es handelt sich um ein stationäres Messsystem mit hoher Messgenauigkeit, mit dem verschiedene physiologische Parameter aufgezeichnet werden können. Je nach Anforderung können folgende Signale aufgezeichnet werden: EKG, PPG, EDA, Atmung und Beschleunigung. Weitere Spezifikationen zu diesem System finden sich bei Gharbi et al. (2008).

Für die im Rahmen dieser Arbeit durchgeführten Studien wurde das Labormesssystem so konfiguriert, dass EKG, PPG und EDA aufgezeichnet werden.



**Abbildung 4.2:** Mind and Body Monitoring Labormesssystem

#### Varioport-e

Bei dem Varioport-e (Becker, 2009) handelt es sich um ein mobiles Aufnahmegerät für EKG, PPG, EDA und Aktivität. Aufgrund der kleinen Größe (55 x 38 x 12 mm) sowie des geringen Gewichtes (ca. 30 g) ist dieses Gerät auch für mobile Aufnahmen geeignet. Die Datenübertragung erfolgt jedoch kabelgebunden. Für Studien in mobilen Szenarien bedeutet dies, dass Daten ex post ausgewertet werden müssen. Online-Übertragung und -Auswertung der Daten an einen PC ist aufgrund des Datenkabels nur in

stationären Szenarien möglich.

### SOMNOscreen

Das SOMNOscreen RC Messsystem von Somnomedics ist ein portables Gerät, welches insbesondere für die Durchführung polysomnographischer Untersuchungen entwickelt wurde. Dieses System ist ebenfalls modular aufgebaut und dadurch problemlos erweiterbar. Auslesen und Auswertung der Messdaten erfolgt offline über eine mitgelieferte Software. Im Rahmen dieser Arbeit wurde das System zur Aufzeichnung von EKG-, EDA- und PPG-Daten verwendet.

### ekgMove

Der von der Firma movisens entwickelte ekgMove-Sensor ist ein mobiles System, das unter der Kleidung mit einem Brustgurt getragen wird. Der Sensor verfügt außerdem über einen 3-Achsen Beschleunigungssensor. Im Gegensatz zum Laborsystem und dem Varioport-e kommen bei diesem System Trockenelektroden zum Einsatz, was auch Aufzeichnungen über einen längeren Zeitraum ermöglicht. Zudem ist der Sensor durch den Brustgurt über einen längeren Zeitraum nahezu belastungsfrei tragbar. Über eine Bluetooth-Schnittstelle können Daten kabellos an einen Computer oder ein Smartphone zur Weiterverarbeitung übertragen werden.



**Abbildung 4.3:** ekgMove-Sensor (movisens GmbH, 2014)

### edaMove

Bei dem ebenfalls von der Firma movisens entwickelten edaMove-Sensor handelt es sich um ein mobiles System zur Aufzeichnung von EDA. Wie

auch der ekgMove-Sensor verfügt der edaMove über einen 3-Achsen Beschleunigungssensor. Der Sensor kann mittels eines Armbands am Handgelenk getragen werden. Die Befestigung der Elektroden erfolgt wahlweise mittels Kleberingen an der Handfläche oder an den Fingern. Die Datenübertragung kann entweder über Bluetooth oder über die USB-Schnittstelle erfolgen.

### 4.2.2 Entwicklung einer Maus zur Erfassung des Tastendrucks

Neben den in Abschnitt 3.3 geschilderten Biosignalen, stellt der Tastendruck auf die Maus ein weiteres Signal dar, das den emotionalen Zustand eines Nutzers reflektiert. Frühere Untersuchungen haben gezeigt, dass die auf eine Maus ausgeübte Kraft zur Unterscheidung verschiedener Nutzerzustände genutzt werden kann. So verwendeten Reynolds & Reynolds (2001) eine mit 8 Kraftsensoren ausgestattete Maus, um zwischen verschiedenen Stufen der Nutzerfrustration zu unterscheiden. In einer anderen Studie zeigten Dennerlein et al. (2003), dass die Kraft, die auf die Seite einer Maus ausgeübt wird, mit höherem Frustrationslevel eines Nutzers ansteigt. Auch in einem von Woolf et al. (2009) entwickelten ITS, das den emotionalen Nutzerzustand berücksichtigt, kam eine drucksensitive Maus zum Einsatz.

Bei den oben angeführten Studien wurden jeweils spezielle Prototypen eingesetzt, die eigens für diese Studien entwickelt wurden. Jedoch sind bisher keine kraftsensitiven Mäuse kommerziell erhältlich. Des Weiteren sind die Sensoren in der Regel an der Außenseite der Maus angebracht, so dass es für einen Benutzer offensichtlich ist, dass die Maus modifiziert wurde. Dies kann unter Umständen dazu führen, dass die auf die Maus ausgeübte Kraft willentlich manipuliert wird, was eine Verfälschung der Messergebnisse bewirkt. Aus diesen Gründen wurde für die vorliegende Arbeit eine Maus entwickelt, in die Kraftsensoren unsichtbar integriert sind. Die folgenden Abschnitte geben einen Überblick über die entwickelte Maus. Für weiterführende Informationen sei auf Degen (2011) verwiesen.

#### 4.2.2.1 Anforderungen an die Maus

Ausgangsbasis für die Entwicklung der kraftsensitiven Maus stellt eine kommerziell erhältliche Computermaus dar. Für die Auswahl der Maus

sowie die Integration der Kraftsensoren in die Maus wurden eine Reihe von Anforderungen definiert, die im Folgenden erläutert werden.

Das Ziel der entwickelten Maus besteht darin, dass diese die auf die Maus ausgeübte Kraft messen soll. Die Kraft soll hierbei an zwei verschiedenen Punkten gemessen werden: die auf die Maustaste ausgeübte Kraft sowie die mit dem Daumen auf die Seite der Maus ausgeübte Kraft. Zusätzlich zu der Kraftmessung müssen die Standard-Mausfunktionalitäten erhalten werden. Die Maus soll in der Lage sein nach Start einer Messung kontinuierlich Daten an einen angeschlossenen PC zu übertragen.

Die wohl wichtigste Anforderung neben der eigentlichen Kraftmessung stellt die unsichtbare Integration der Kraftsensoren dar. Diese sollen so integriert werden, dass sie für einen Benutzer nicht ersichtlich sind. Dies bedeutet, dass die zu modifizierende kommerzielle Computermaus ausreichend Platz bieten muss, um eine zusätzliche Platine zu integrieren.

Für die an den beiden Messpunkten integrierten Kraftsensoren gelten unterschiedliche Anforderungen:

- *Kraftaufnahme unter der Maustaste:* Aufgrund der sehr feinen Kraftabstufungen, die durch den menschlichen Finger ausgeübt werden können, ist es wichtig, dass der verwendete Kraftsensor über eine sehr hohe Auflösung verfügt.
- *Kraftaufnahme an der Mausseite:* Der hierfür verwendete Kraftsensor muss flach genug sein, damit er problemlos in der Maus platziert werden kann.

Des Weiteren wird ein Prozessor mit einem A/D-Wandler benötigt, der die kontinuierlichen Eingangssignale des Sensors digitalisiert und an einen Computer sendet. Hierzu soll der USB-Anschluss der Maus verwendet werden um zusätzliche Kabel zu vermeiden. Der Stromanschluss für die analogen und digitalen Schaltkreise soll onBoard realisiert werden. Die Erfassung der Sensordaten soll mit einer festgelegten Samplingfrequenz erfolgen, die hoch genug ist, um eine exakte Analyse der Daten zu ermöglichen.

### 4.2.2.2 Konzeption

Nach der Analyse verschiedener kommerzieller Computermäuse wurde die Maus *Revoltec W102* aus dem Gaming-Bereich ausgewählt. Der Vorteil



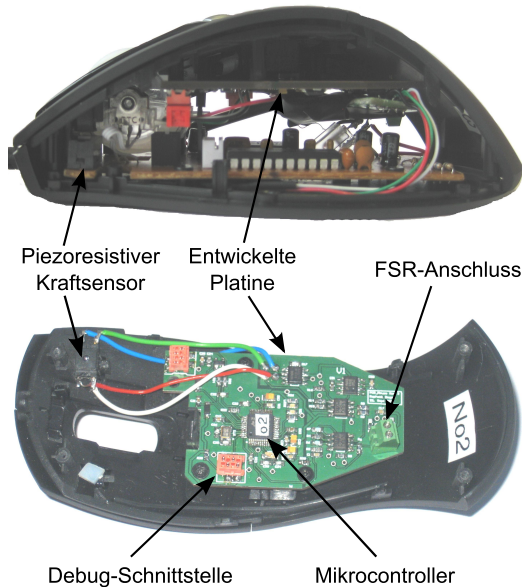
dieser Maus besteht darin, dass diese ausreichend Platz zur unsichtbaren Integration zusätzlicher Komponenten in die Maus bietet. Zur Messung der Kraft auf die Maustaste wurde der piezoresistive Sensor Honeywell FSS1500 ausgewählt. Dieser ermöglicht eine sehr präzise Erfassung von Kräften zwischen 0 und 1500 N, was deutlich über der Kraft liegt, die unter normalen Bedingungen auf eine Maustaste ausgeübt wird.

Nutzt man diesen Sensor in einer Wheatstone-Brücke, wird der Temperaturdrift piezoresistiver Materialien größtenteils ausgeglichen. Der Instrumentenverstärker AD627 kann verwendet werden, um die Ausgangsspannung der Wheatstone-Brücke auf die Größe des A/D-Wandler zu verstärken.

Um den seitlichen Druck auf die Maus zu messen, wurde der kraftsensitive Widerstand Interlink Electronics FSR400 ausgewählt. Der Widerstandswert wird mit Hilfe eines Strom-Spannungswandlers in eine Spannung umgewandelt, so dass der A/D-Wandler den zugrundeliegenden Widerstand basierend auf der Kraft, die auf den Sensor ausgeübt wird, messen kann. Für die Integration des Systems wurde der Microcontroller MSP430F5510 ausgewählt. Der sehr kleine Mikrocontroller (9x9 mm) ist in der Lage, die analogen Signale zu messen und verfügt des Weiteren über eine USB-Schnittstelle. Durch einen eingebauten 3,3 V Low Drop-out Spannungsregler sind für diesen Chip nur wenige zusätzliche Bauteile erforderlich. Aufgrund des geringen Stromverbrauchs ist es möglich, den Mikrocontroller über die USB-Schnittstelle mit Strom zu versorgen. Die Schnittstelle zwischen der Maus und dem Computer wird mittels eines USB-CDC realisiert. In dieser USB-Klasse können die Windows-spezifischen Maustreiber verwendet werden. Es wird lediglich eine zusätzliche Initialisierungsdatei benötigt, um die modifizierte Maus zu installieren. Zusätzlich wird ein USB-Hub auf der Platine integriert, um die gleichzeitige Kommunikation mit der Kraftmessung und den Standard-Mausfunktionalitäten zu ermöglichen.

### 4.2.2.3 Integration der Sensorik in die Maus

Die Revoltec W102-Maus enthält zwei Platinen. Während die erste Platine die Hauptfunktionalitäten der Maus wie Klicken und Mausbewegungen übernimmt, stellt die zweite Platine die Funktionalitäten für weitere Mausbuttons – wie beispielsweise vorwärts und rückwärts zu blättern – zur



**Abbildung 4.4:** Integration der Sensorik in die Maus

Verfügung. Da diese Funktionalitäten in den Zielanwendungen nicht benötigt werden, wurde diese Platine durch die neu entwickelte Platine zur Kraftmessung ersetzt. Hierbei war es wichtig, dass das Platinenlayout unverändert blieb.

Neben dem Mikrocontroller mussten auch die Anlogschaltung, die Anschlüsse für die Sensoren und das Debugging sowie der USB-Anschluss auf dem Board platziert werden. Abbildung 4.4 stellt dar, wie die zusätzlichen Komponenten in die Maus integriert wurden.

Der Mikrocontroller digitalisiert die Daten von den Kraftsensoren mit einer Samplingfrequenz von 32 Hz und sendet diese blockweise mit einer Samplingfrequenz von 1 Hz an einen Computer. Auf dem Computer muss eine Initialisierungsdatei erstellt werden, die die Informationen über die CDC USB-Quelle (z.B. Taktfrequenz oder Hersteller ID) enthält. Bei Verbindung der Maus mit einem Windows-Computer, erkennt das Betriebssystem ein USB-Hub mit zwei verbundenen Geräten: eine Human

Interface Device und ein CDC-Gerät. Die Installation der HID-Klasse wird automatisch von Windows vorgenommen. Die Installation des virtuellen COM-Ports muss manuell mittels einer Installationsdatei durchgeführt werden. Mithilfe der Installationsdatei generiert Windows einen virtuellen COM-Port zu der Maus. Dieser COM-Port ermöglicht die bidirektionale Kommunikation zwischen dem Computer und der modifizierten Maus.

### 4.2.3 Einordnung der verwendeten Messsysteme

Für die im Rahmen dieser Arbeit durchgeführten Studien waren unterschiedliche Kriterien ausschlaggebend für die Wahl der Systeme zur Aufzeichnung der physiologischen Daten. Für die Laborstudien zur Evaluation der physiologischen Parameter war bei der Sensorauswahl in erster Linie die Verfügbarkeit einer Vielzahl zeitsynchron aufgezeichneter Signale ausschlaggebend. Bei der Konzeption der Lernanwendungen hingegen, spielte die Mobilität der Sensoren eine bedeutende Rolle. Tabelle 4.1 zeigt, wie die ausgewählten Geräte zur Aufzeichnung der ausgewählten Biosignale gemäß der in Abschnitt 3.1 dargestellten Kriterien eingeordnet werden können.

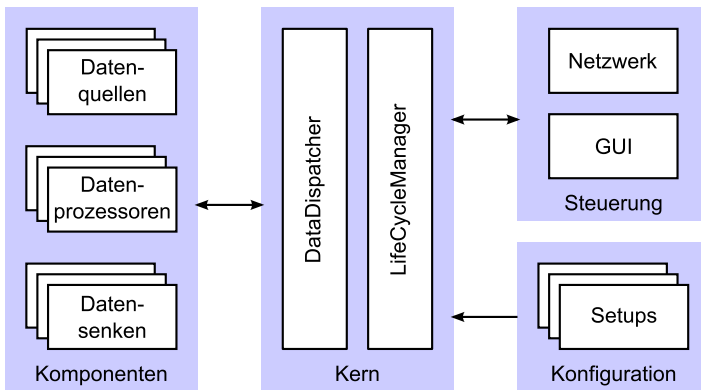
	Druckmaus	Varioport-e	SOMNOscreen	Labormesssystem	ekgMove	edaMove
Mobilität	-	o	+	-	+	+
Onlinefähigkeit	+	+	-	+	+	+
Kommerzielle Verfügbarkeit	-	+	+	o	+	+
Medizinproduktzulassung	-	-	+	-	-	-
Offenheit	+	+	-	+	+	+
Erweiterbarkeit	-	-	+	+	-	-

**Tabelle 4.1:** Einordnung der ausgewählten Messsysteme

## 4.3 xAffect – Ein Software-Framework für die Online-Analyse physiologischer Daten

Zur Realisierung der Datenverarbeitungsschicht wurde mit xAffect eine Middleware entwickelt, welche ein modulares Framework für Anwendungen zur Verfügung stellt, die mit geringer Latenz Feedback aus physiologischen Daten integrieren.

Im Zentrum bei der Entwicklung der xAffect-Software stand die Anforderung nach Modularität und Konfigurierbarkeit. xAffect wurde daher so konzipiert, dass eine Vielzahl unterschiedlicher Funktionalitäten miteinander verknüpft werden können. Hierzu wurde die Funktionalität der Anwendung in einzelne modulare Komponenten ausgelagert, welche flexibel miteinander kombiniert werden können. Abbildung 4.5 stellt den Aufbau der Software dar.



**Abbildung 4.5:** Gesamtarchitektur der xAffect-Software

Die Komponenten können in drei verschiedene Kategorien untergliedert werden:

*Datenquellen* beinhalten alle Arten eingehender Daten wie beispielsweise Sensordaten, Signalgeneratoren oder Timer. Sie können anderen Komponenten Daten zur Verfügung stellen, sind jedoch nicht in der Lage, Daten zu empfangen.

*Datenprozessoren* sind verantwortlich für die Verarbeitung von Daten. Sie kombinieren die Eigenschaften von Datenquellen und Datensenzen, da sie in der Lage sind, sowohl Daten zu empfangen als auch zu senden.

*Datensenzen* stellen das Gegenstück zu den Datenquellen dar und können für das Logging und Übermitteln von Daten an andere Anwendungen verwendet werden. Im Gegensatz zu Datenquellen können sie ausschließlich Daten empfangen jedoch keine Daten an andere Komponenten senden.

Die Komponenten können über sogenannte Setups flexibel miteinander kombiniert werden. In Abschnitt 4.3.1 wird hierauf genauer eingegangen. Die Verknüpfung der Komponenten erfolgt durch den `DataDispatcher`. Die Komponenten in `xAffect` unterliegen einem klar definierten Lebenszyklus. Der `LifeCycleManager` hat hierbei die Aufgabe, jederzeit einen konsistenten Systemzustand zu gewährleisten. Der Lebenszyklus der Komponenten wird in Abschnitt 4.3.2 genauer erläutert.

Die Software ist aus zwei separaten Bestandteilen aufgebaut: einer plattformunabhängigen Bibliothek, welche die grundlegenden Funktionalitäten der Software implementiert, sowie der `xAffect` Desktop-Applikation, welche es ermöglicht, `xAffect` als eigenständige Anwendung zu nutzen. Durch diese Trennung wird sowohl die Anforderung nach Plattformunabhängigkeit erfüllt als auch die Verwendung als Standalone-Anwendung ermöglicht. Im Gegensatz zu der `xAffect`-Bibliothek ist die Desktop-Applikation nicht plattformunabhängig, da diese auf dem Standard-Widget-Toolkit (SWT) basiert. Aufgrund der umfangreichen Darstellungsmöglichkeiten, die von SWT insbesondere für die Online-Visualisierung geboten werden, und da die `xAffect`-Software in der aktuellen Arbeit nur auf Windows-Systemen verwendet wird, wurde die Anforderung der Plattformunabhängigkeit für die Darstellungsebene vernachlässigt. Jedoch ist die Implementierung flexibel genug gestaltet, dass die Windows-Version durch Ersetzen des SWT durch andere Bibliotheken zur Darstellung auch auf andere Plattformen portiert werden kann.

Die `xAffect`-Bibliothek stellt die Grundfunktionalitäten der Software zur Verfügung. Hierzu gehören die modularen abstrakten Datenstrukturen, welche genutzt werden können, um das System durch zusätzliche Anwendungslogik oder Algorithmen zu erweitern. Diese wurden aus dem

Unisens-Datenformat abgeleitet. Zudem stellt die Bibliothek Mechanismen zur Konfiguration von Setups sowie dem Lebenszyklus-Management zur Verfügung. Ein weiterer wichtiger Bestandteil sind Schnittstellen zur Kontrolle und Steuerung der Software – je nach der spezifischen Anforderung aus der Anwendung kann dies über die Desktop-Applikation oder aber über das Netzwerk erfolgen, wenn die xAffect-Software in komplexe Studienszenarien mit 3rd-Party Software eingebunden wird.

Die xAffect Desktop-Applikation erweitert die Bibliothek um die notwendigen Funktionalitäten, um die Software als eigenständige Anwendung nutzen zu können. Zum einen wird hierfür eine GUI zur Steuerung der Software zur Verfügung gestellt, zum anderen werden Möglichkeiten zur Datenvisualisierung angeboten.

### 4.3.1 Verknüpfung von Komponenten

Zur Gewährleistung der Modularität ist die Anwendungslogik von xAffect in modularen Komponenten gekapselt. Diese können gemäß der jeweiligen Anforderungen einer Studie verknüpft werden. Dies geschieht über die Nutzung von Setups über welche die Kombination und Konfiguration der für eine Studie benötigten Komponenten erfolgt.

Szenarien, bei welchen mehrere Sensoren und Algorithmen kombiniert werden sollen, führen oftmals zu komplexen Datenflussmustern. Durch Herunterbrechen auf vier elementare Datenflussmuster, welche beliebig kombiniert werden können, sind auch komplexe Datenflüsse problemlos mit xAffect modellierbar. In Abbildung 4.6 sind diese elementaren Muster dargestellt.

Das Einfachste dieser Muster (a) besteht aus einer linearen Aneinanderreihung der Komponenten, wobei die Ausgabe einer Komponente die Eingabe für die nächste Komponente darstellt.

Ein weiteres Datenflussmuster stellt die Parallelverarbeitung (b) dar. Hierbei werden die Ausgaben einer Komponente an zwei oder mehr weitere Komponenten weitergegeben. Dies ermöglicht beispielsweise eine Beschleunigung bei der echtzeitnahen Datenverarbeitung.

Insbesondere in multimodalen Anwendungsszenarien, bei welchen mehr als ein Sensor zur Aufzeichnung von Daten verwendet wird, kann es vorkommen, dass Informationen von unterschiedlichen Signalen fusioniert

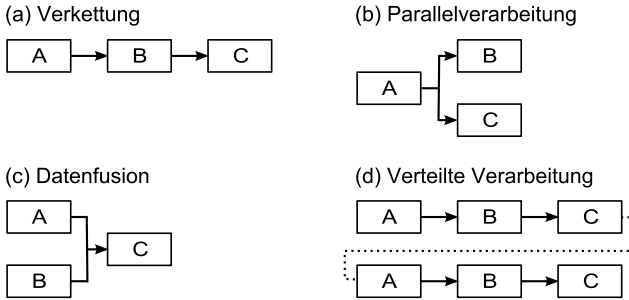


Abbildung 4.6: Elementare Datenflussmuster

werden müssen. Auch die Datenfusion (c) mehrerer Signale – auch mit unterschiedlichen Samplingraten – ist mit xAffect problemlos möglich.

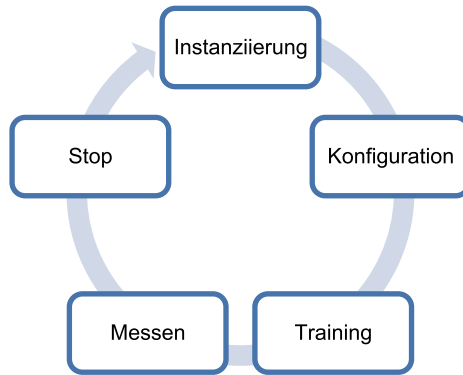
Zuletzt ermöglicht xAffect auch eine verteilte Datenverarbeitung (d) was insbesondere für mobile oder cloudbasierte Anwendungen von Bedeutung ist.

Die dargestellten Datenflussmuster können beliebig miteinander kombiniert werden, so dass auch komplexe Datenflüsse problemlos mithilfe der von xAffect zur Verfügung gestellten Strukturen erstellt werden können. Zu beachten ist hierbei jedoch, dass die Schnittstellen der Ausgangskomponenten mit denen der Eingangskomponenten kompatibel sein müssen.

### 4.3.2 Lebenszyklus der Komponenten

Wie einführend erläutert, unterliegen die Komponenten einem klar definierten Lebenszyklus. Um einen konsistenten Systemzustand zu gewährleisten, befinden sich alle an einer Prozesskette beteiligten Komponenten immer am selben Punkt ihres Lebenszyklusses. Der Lebenszyklus der Komponenten in xAffect ist in Abbildung 4.7 veranschaulicht.

Das Management des Lebenszyklus innerhalb der Software erfolgt durch den LifeCycleManager. Die Phasen des Lebenszyklus orientieren sich an den Anforderungen, die sich aus dem typischen Ablauf psychophysiologischer Studien ableiten und werden im Weiteren näher erläutert.



**Abbildung 4.7:** Lebenszyklus der Komponenten

**Instanziierung** Während der Instanziierung werden die für eine Studie benötigten Komponenten definiert und verknüpft. Durch Aufrufen des Konstruktors einer Komponente, registriert sich diese automatisch beim LifeCycleManager. Zudem führt der LifeCycleManager eine Verknüpfung der Komponenten gemäß der angegebenen Spezifikation durch und gleicht die Ein- und Ausgabeparameter der jeweiligen Komponenten miteinander ab. Die anschließende Parametrisierung erfolgt basierend auf der in der Konfigurationsstruktur einer Komponente vorgegebene Informationen, welche alle für die Instanziierung der jeweiligen Komponenten erforderlich Informationen erhält. Zuletzt erfolgt eine Überprüfung der zuvor spezifizierten Parameter auf Gültigkeit. Sind einzelne Parameter falsch konfiguriert, wird eine Fehlermeldung zurückgegeben.

**Konfiguration** Im Anschluss an die Instanziierung der Komponenten erfolgt deren Konfiguration. Während dieser Phase bereiten sich die Komponenten entsprechend der im Konstruktor festgelegten Konfiguration auf das Senden und Empfangen der hier festgelegten Daten vor. Die Konfigurationsphase beinhaltet somit Prozesse wie die Verbindungsherstellung zu Sensoren sowie das Anlegen von Logfiles.



**Training** Die Trainingsphase dient der Initialisierung der in einem Setup verwendeten Komponenten. Während des Trainings können die Komponenten die eingehenden Daten zur Kalibrierung der späteren Messung nutzen. So erfordern viele Studienszenarien vor der eigentlichen Messung die Aufzeichnung einer Baseline-Phase, welche als Referenz für die späteren Messdaten verwendet werden kann. Mit Beginn der Trainingsphase starten alle Komponenten mit dem Senden der Daten, so dass diese auch für andere Komponenten zur Verfügung stehen. Für den Fall, dass einzelne Komponenten für das Training eine bestimmte Mindestdauer erfordern, besteht die Möglichkeit, dass der Beginn der Messung durch andere Komponenten geblockt wird. Nachdem ausreichend Trainingsdaten gesammelt sind, kann der Start der Messung wieder aktiviert werden.

**Messen** Die Messphase stellt die wichtigste Phase im Lebenszyklus der Komponenten dar. In dieser Phase produzieren die zuvor in der Trainingsphase initialisierten Komponenten die Nutzdaten für die jeweiligen Anwendungen. Je nach Anforderungen können diese Messdaten gespeichert und / oder an eine andere Anwendung weitergegeben werden.

**Stop** Beim Beenden einer Messung erfolgt eine Benachrichtigung aller für die Messung aktiven Komponenten. Dies bewirkt, dass temporäre Datenstrukturen verworfen sowie Verbindungen zu anderen Komponenten wieder freigegeben werden. Nach Beenden der Messung kann die xAffect-Software zurückgesetzt und der Lebenszyklus der Komponenten erneut initiiert werden.

### 4.3.3 Datenstrukturen

Die internen Datenstrukturen von xAffect wurden in Anlehnung an das Unisens-Datenformat entwickelt. Die Beschreibung von eingehenden und ausgehenden Daten erfolgt mittels der DataDescription. Diese beinhaltet folgende Elemente:

- *Content Class*: gibt die Art der Daten an
- *Anzahl der Kanäle*: spezifiziert die Anzahl der Datenkanäle
- *Samplingfrequenz*: Abtastrate der Daten in Hertz

- *Datentyp*: Typ der Daten. Wie auch bei Unisens wird hier zwischen DataEvents, DataSignals und DataValues unterschieden.

Sämtliche aufgezeichneten und verarbeiteten Daten können im Unisens-Datenformat gespeichert werden. Dies vereinfacht die spätere Offline-Analyse mit bereits für die Verarbeitung von Unisens-Datensätzen zur Verfügung stehenden Tools.

### 4.3.4 Steuerung der xAffect-Software

Zur Kontrolle der xAffect-Software gibt es zwei verschiedene Möglichkeiten: Zum einen kann die Software über via UDP über die Netzwerkschnittstelle gesteuert werden, zum anderen über eine GUI.

Bei Nutzung von xAffect über die Netzwerkschnittstelle erfolgt Konfiguration und Steuerung der Software über ein XML-basiertes Protokoll über welches der komplette Lebenszyklus der Software gesteuert werden kann. Listing 4.1 zeigt beispielhaft eine Konfigurationsnachricht, welche von einer externen Anwendung beim Start von xAffect gesendet wird.

**Listing 4.1:** Beispiel einer Konfigurationsnachricht

```
<config>
  <subjectInfo subjectId="03" sessionId="04"
    institutionId="FZI" studyId="S-T01" />
  <dataSetup name="org.xaffect.setup.nBack">
    <parameter key="age" value="25"/>
    <parameter key="sex" value="female"/>
  </dataSetup>
</config>
```

Die Nutzeroberfläche von xAffect ist in Abbildung 4.8 dargestellt. Die Oberfläche wurde an die Anforderungen physiologischer Studien angepasst. So können im oberen Teil der Oberfläche die Informationen über die Versuchspersonen und die aktuelle Studie angegeben werden. Die Konfiguration ermöglicht die Auswahl eines vordefinierten Setups, welches die benötigten Komponenten verbindet. Mittels des 'Configure'-Buttons wird das Setup an die xAffect-Software übermittelt. Zusätzlich können die einzelnen Phasen des Experiments über die Oberfläche gesteuert werden. Ein Logfenster gibt Auskunft über den jeweiligen Zustand der Software.

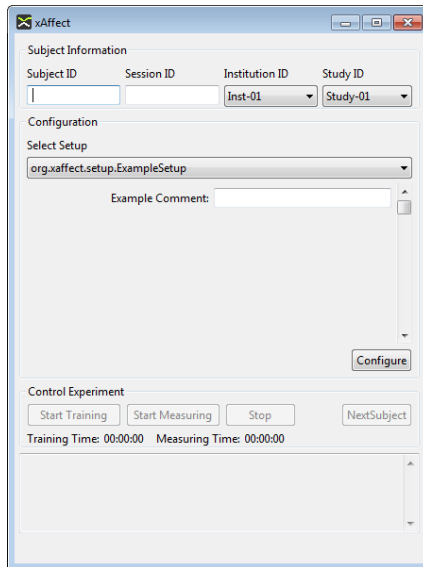


Abbildung 4.8: Nutzeroberfläche der xAffect-Software

### 4.3.5 Integrierte Komponenten

Im Folgenden werden die Komponenten beschrieben, die im Rahmen dieser Arbeit für die xAffect-Software entwickelt wurden. Die Darstellung erfolgt in Anlehnung an die xAffect-Struktur aufgeteilt in Datenquellen, Datenprozessoren und Datensenken.

#### Datenquellen

**ekgMove-Adapter** Der ekgMove-Adapter bindet den ekgMove-Sensors an xAffect an. Hierfür wurde das mitgelieferte Java SDK verwendet, welches alle benötigten Funktionen hinsichtlich Konnektivität und Datenaustausch zur Verfügung stellt. Die Kommunikation zwischen dem Sensor und xAffect erfolgt kabellos über Bluetooth.

**Labormesssystem-Adapter** Die Anbindung des Labormesssystems erfolgt über eine Labview-Runtime-Umgebung, welche einen lokalen Server zur Verfügung stellt. xAffect nimmt hierbei die Rolle des Clients ein. Die Kommunikation zwischen den beiden Systemen erfolgt über UDP. Entsprechend der Studienanforderungen kann der Adapter derart konfiguriert werden, dass alle benötigten Signale von xAffect erfasst werden.

**Mausadapter** Die Integration der kraftsensitiven Maus erfolgt analog zum Varioport-e mittels der RXTX-Bibliotheken. Die Übertragung der Daten erfolgt paketweise mit einer Samplerate von 32 Hz über den USB-Anschluss.

**Timer** Der Timer dient dem regelmäßigen Triggern von Ereignissen. Als Ausgangssignal wird ein DataEvent mit einer frei parametrierbaren Frequenz gesendet.

**Unisens-Reader** Mithilfe des Unisens-Readers können zuvor aufgezeichnete Unisens-Datensätze eingelesen werden. Diese Komponente wurde im Rahmen dieser Arbeit zum Testen anderer Komponenten verwendet, fand in den durchgeführten Studien jedoch keinen Einsatz.

**Varioport-e-Adapter** Die Kommunikation zwischen der xAffect-Software und dem Varioport-e erfolgt kabelgebunden über eine serielle RS232-Schnittstelle. Zur Anbindung des Sensors an die xAffect-Software wurden die RXTX-Bibliotheken verwendet. Je nach benötigter Konfiguration kann auch bei diesem Adapter flexibel eingestellt werden, welche Signale seitens xAffect verarbeitet werden.

### Datenprozessoren

**QRS-Detektion** Die QRS-Detektion wird mittels des OSEA-Algorithmus (Hamilton, 2002) vorgenommen. Als Eingangssignal wird ein EKG-Signal erwartet. Der Prozessor liefert als Ausgabe Events, welche die Zeitpunkte der detektierten QRS-Komplexe enthalten.

**HR-Berechnung** Die Komponente zur Berechnung der Herzrate verwendet als Eingang die mittels der Komponente zur QRS-Detektion identifizier-

ten Triggerzeitpunkte. Basierend auf diesen wird die aktuelle Herzrate berechnet und ausgegeben.

**Erregungs-Berechnung** Der Datenprozessor zur Errechnung des Erregungsniveaus bestimmt die physiologische Erregung basierend auf der Herzrate. Als Eingabe werden die Werte aus der Herzratenberechnung sowie der Timer verwendet. Die Berechnung der Erregung erfolgt relativ zu der während des Trainings aufgezeichneten Baseline-Herzrate. Die Spannweite des Erregungsniveaus sowie die Fensterbreite zur Berechnung der Erregung ist bei dieser Komponente parametrierbar gehalten. Das berechnete Erregungsniveau wird mit der durch den Timer vorgegebenen Samplingrate ausgegeben.

**EKG-Artefakt-Detektion** Die Komponente zur Artefakt-Detektion dient der Annotation fehlerhaft detektierter R-Zacken. Als Eingabe dienen hierbei die Trigger aus der QRS-Detektion. Zur Erkennung der Artefakte wird das von Clifford (2007) beschriebene Verfahren verwendet. Fehlerhaft erkannte Trigger werden mit *Q* (*unclassifiable beat*), korrekte Trigger mit *N* (*normal beat*) annotiert.

**EKG-Feature-Extraktion** Basierend auf den artefaktbereinigten Triggern berechnet diese Komponente Herzrate und RMSSD einer Messung auf einem Zeitfenster konfigurierbarer Länge. Die Berechnung kann durch eine externe Anwendung über das Netzwerk gestartet werden.

### Datensenken

**Unisens-Logging** Der Unisens-Datenlogger dient der Speicherung der erfassten und berechneten Daten. Neben den erfassten Rohdaten können auch sämtliche hieraus abgeleiteten Daten mittels dieser Komponente im Unisens-Datenformat gespeichert werden. Die Anzahl der geloggteten Unisens-Entries ist hierbei nicht begrenzt.

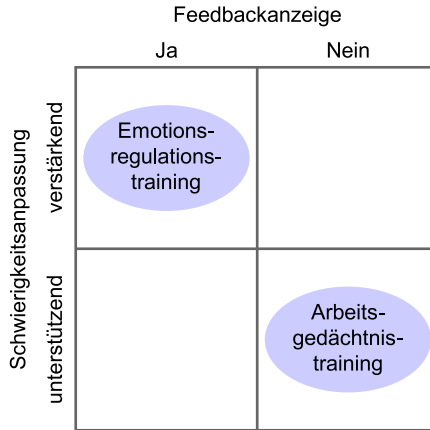
**Unisens-Streaming** Diese Komponente dient dem Streaming von Daten an andere Anwendungen. Die Übertragung erfolgt mittels UDP. Das zur Datenübertragung genutzte Protokoll orientiert sich am Unisens-Datenformat.

## 4.4 Lernanwendungen zur Interaktion mit dem Nutzer

Im Rahmen dieser Arbeit werden zwei unterschiedliche Lernanwendungen vorgestellt, welche darauf abzielen, den einleitend vorgestellten Regelkreis des technologiegestützten Lernens zu verbessern: Auf der Nutzerseite wird ein System entwickelt, welches diesem mittels physiologischem Feedback helfen soll, seinen eigenen emotionalen Zustand an das Lernsystem anzupassen. Auf der Systemseite steht die Adaption des Systems an den emotionalen Zustand des Lernenden mittels physiologischem Feedback im Vordergrund.

Grundsätzlich gibt es bei der Integration von Informationen über die emotionale Erregung in Lernanwendungen verschiedene Möglichkeiten. Die beiden wichtigsten Dimensionen stellen hierbei die Präsentation des Feedbacks sowie die Art der Anpassung der Lernschwierigkeit an Veränderungen des emotionalen Erregungsniveaus dar (siehe Abbildung 4.9). Je nach Art der Anwendung kann die Anzeige des Feedbacks über den Nutzerzustand Vor- und Nachteile mit sich bringen. Ebenso kann es je nach Anwendungsgebiet sinnvoll sein, den Schwierigkeitsgrad in Abhängigkeit des emotionalen Erregungsniveaus eines Nutzers unterstützend oder verstärkend anzupassen. Bei der unterstützenden Anpassung des Schwierigkeitsgrades reagiert die Lernanwendung auf Abweichungen vom optimalen Erregungsniveau derart, dass die Schwierigkeit dahingehend angepasst wird, dass das optimale Erregungsniveau möglichst schnell wieder erreicht wird. Bei der verstärkenden Anpassung hingegen wird eine Abweichung vom optimalen Erregungsniveau gewissermaßen bestraft, indem die Anwendung diese Abweichungen durch eine Veränderung der Trainingsschwierigkeit weiter verstärkt.

Die im Rahmen dieser Arbeit vorgestellten Lernanwendungen decken beide Varianten der Feedbackpräsentation sowie der Schwierigkeitsanpassung ab. Beim Emotionsregulationstraining wird erfolgreiche Regulation der emotionalen Erregung mit einer leichteren Trainingsschwierigkeit belohnt. Die Feedbackanzeige ist bei diesem Training essentiell, da der Nutzer diese Information benötigt, um seinen Erregungszustand entsprechend zu regulieren. Beim Arbeitsgedächtnistraining hingegen steht bei der Adaption das Ziel im Vordergrund, den Nutzer zu jedem Zeitpunkt optimal zu fordern. Weicht das Erregungsniveau vom optimalen Erregungsniveau ab,



**Abbildung 4.9:** Varianten zur Integration von physiologischem Feedback in Lernanwendungen

erfolgt eine Anpassung der Trainingsschwierigkeit derart, dass der Nutzer möglichst schnell wieder einen optimalen emotionalen Erregungszustand erreicht. Im Gegensatz zum Emotionsregulationstraining kann die Anzeige des Nutzerzustands beim Arbeitsgedächtnistraining den Trainingsverlauf negativ beeinflussen, da dies von der eigentlichen Aufgabe ablenkt. Aus diesem Grund erfolgt hier keine Feedbackanzeige.

Die beiden vorgestellten Anwendungen veranschaulichen die Vielschichtigkeit bei der Integration von emotionaler Erregung in Lernumgebungen. Die für eine Anwendung erforderlichen Adaptionenmechanismen hängen somit sehr individuell von den Anforderungen einer Lernanwendung ab.

# 5 Erkennung von Erregung aus physiologischen Signalen

Bevor eine Integration von physiologischem Feedback in eine Lernumgebung erfolgen kann, muss zunächst untersucht werden, in welchen physiologischen Signalen sich emotionale Erregung am besten widerspiegelt. Es wurde daher eine Studie mit dem Ziel durchgeführt, die physiologischen Parameter bzw. Kombination physiologischer Parameter zu bestimmen, welche den emotionalen Zustand einer Person am besten widerspiegeln. Zwar gibt es hierzu bereits einige Anhaltspunkte in der Literatur, jedoch wurde bisher keine personenunabhängige Studie mit der Kombination aus EKG, PPG, EDA sowie Maus-Kraftdaten durchgeführt, was die nachfolgend dargestellten Untersuchungen erforderlich machte.

In Abschnitt 5.1 erfolgt zunächst eine Einführung in das verwendete Studiendesign. Im Anschluss daran ist die Analyse der gesammelten Daten sowie der Aufbau des multimodalen Klassifikators beschrieben. Abschließend erfolgt eine Zusammenfassung der wichtigsten Erkenntnisse aus dieser Studie.

Teile dieses Kapitels wurden bereits in abgeänderter Form in Schaaff, Degen et al. (2012); Schaaff & Adam (2013) veröffentlicht.

## 5.1 Studiendesign

Das Ziel dieser Studie besteht darin, herauszufinden, welche physiologischen Signale am besten mit dem emotionalen Erregungsniveau einer Person korreliert sind. Das Messen emotionsbezogener Zustände setzt voraus, dass die Personen, deren Signale gemessen werden, sich in einem emotionalen Zustand befinden. Zur Induktion der emotionalen Erregung wurde ein Spiel eingesetzt. Aufgrund der besseren Möglichkeiten zur Kontrolle der Umgebungsvariablen wurde die Studie als Laborstudie durchgeführt.



Das Studiendesign sowie der Ablauf bei der Studiendurchführung sind nachfolgend beschrieben.

### 5.1.1 Studienaufbau

Während der Studie kamen unterschiedliche Hard- und Softwarelösungen zum Einsatz, welche in den folgenden Abschnitten kurz dargestellt werden.

#### 5.1.1.1 Hardware

Im Rahmen der Studie erfolgte eine kontinuierliche Erfassung verschiedener Biosignale. Folgende Signale wurden aufgezeichnet:

- Herzaktivität mittels EKG- und PPG-Sensoren
- Hautleitwert mittels EDA-Sensorik
- Kraftauswirkung auf die Maus

Da die Studie im Labor durchgeführt wurde, wurde das Kriterium der Mobilität bei der Auswahl der Messsysteme für diese Studie nicht in Betracht gezogen. Die Kraftauswirkung auf die Maus wurde mit Hilfe der in Abschnitt 4.2.2 dargestellten modifizierten kraftsensitiven Maus erfasst. Für die Erfassung der weiteren Signale wurden zwei verschiedene Messsysteme verwendet: das Labormesssystem sowie der Varioport-e (siehe Abschnitt 4.2.1). Beide Systeme wurden so konfiguriert, dass jeweils Herzaktivität und Hautleitwert von einem System aufgezeichnet wurden.

#### 5.1.1.2 Software

Bei Studien zur Untersuchung psychophysiologischer Reaktionen auf Emotionen ist die Induktion der gewünschten Emotionen ein wesentlicher Bestandteil der Studie. Aus diesem Grund wurde im Rahmen dieser Studie ein Spiel entwickelt, welches der gezielten Induktion emotionaler Erregung dient. Das Spiel steht in zwei unterschiedlichen Varianten zur Verfügung: Die erste Variante dient der Erzeugung eines möglichst hohen Erregungsniveaus (*high arousal*, HA) bei den Versuchspersonen. Bei der zweiten Variante hingegen wurde auf alle erregungserzeugenden Elemente verzichtet (*low arousal*, LA), so dass diese Variante für den Einsatz bei einer Kontrollgruppe geeignet ist.

Die Aufgabe bei dem Spiel besteht in der Wiedererkennung einer vorgegebenen Kombination von fünf Symbolen, jedoch sind die Aufgaben je nach Spielvariante unterschiedlich schwierig gestaltet. Die beiden Varianten des Spiels sind nachfolgend dargestellt.

### Induktion hoher emotionaler Erregung

Für die Induktion hoher emotionaler Erregung wurde eine Vielzahl erregungserzeugender Elemente in das Spiel integriert. Gegenüber der anderen Variante des Spiels wurde der Schwierigkeitsgrad erhöht, indem die vorgegebene Symbolkombination unter 20 verschiedenen Symbolreihenfolgen identifiziert werden muss. Zusätzlich wird hierbei Zeitdruck erzeugt, da nur 7 Sekunden für die Entscheidung zur Verfügung stehen. Je nach Ergebnis werden folgende Punktzahlen vergeben:

- **20 Pluspunkte** für jede richtig erkannte Symbolreihenfolge
- **10 Minuspunkte** für jede falsch ausgewählte Symbolreihenfolge
- **30 Minuspunkte** wenn innerhalb der 7 Sekunden keine Auswahl getroffen wurde

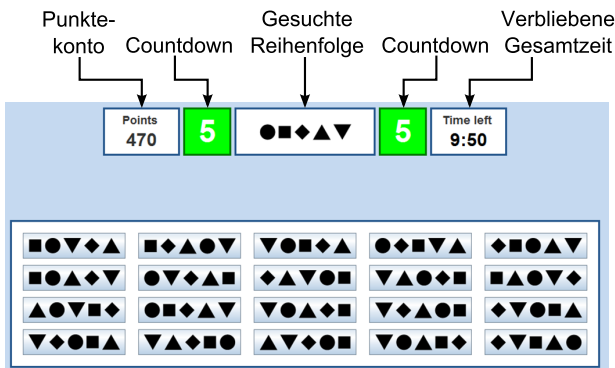


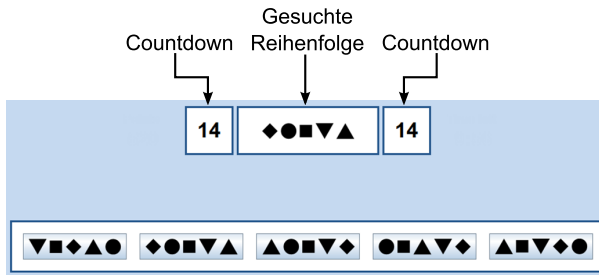
Abbildung 5.1: Spieldesign zur Induktion hoher emotionaler Erregung

Neben dem Countdown für die verbleibende Zeit der jeweiligen Entscheidung wird zusätzlich der aktuelle Punktestand sowie die verbleibende Gesamtzeit angezeigt. Der Bildschirm, den die Spieler während des Spiels präsentiert bekommen, ist in Abbildung 5.1 dargestellt.

Zusätzlich zu den grafischen Elementen wird die Hintergrundmusik im Verlauf des Spieles immer weiter gesteigert, um hierdurch die Erregung zusätzlich zu verstärken. In der letzten Spielminute wird die Melodie noch einmal komplett geändert, um die letzte Spielphase kenntlich zu machen.

### Induktion niedriger emotionaler Erregung

Bei dem zur Induktion niedriger Erregung verwendeten Spiel wurden alle erregungsinduzierenden Elemente aus dem Spiel entfernt. Im Gegensatz zu der erregungsinduzierenden Version des Spiels muss der Spieler die Symbolreihenfolge zwischen lediglich fünf verschiedenen Kombinationen wiederfinden. Für jede Entscheidung wurden insgesamt 14 Sekunden Zeit eingeräumt, so dass kein Zeitdruck entsteht. Der Spielbau ist in Abbildung 5.2 dargestellt.



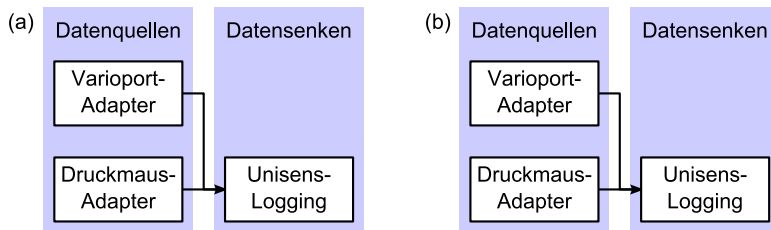
**Abbildung 5.2:** Spieldesign zur Induktion niedriger emotionaler Erregung

Um die emotionale Erregung zusätzlich zu steigern, wurde in dem HA-Szenario zusätzlich die Auszahlung am Ende des Experiments an die Performance während des Spiels geknüpft. Die Höhe der Auszahlung richtete sich nach dem Rang, den man im Verhältnis zu den anderen fünf Spielern erzielt hatte. Die Höhe der Auszahlung ist in Tabelle 5.1 dargestellt. In der Gruppe mit niedriger emotionaler Erregung hingegen, hatte die Performance während der Aufgabe keinen weiteren Einfluss.

Zur zeitsynchronen Aufzeichnung der physiologischen Signale wurde die xAffect-Software mit der in Abbildung 5.3 dargestellten Konfiguration verwendet. Da in diesem Setup keine Rückkopplung der Sensordaten mit dem Spiel vorhanden ist, werden keine Datenprozessoren verwendet.

Rang	Auszahlung [€]
1	15,00
2	12,00
3	9,00
4	6,00
5	3,00
6	0,00

**Tabelle 5.1:** Auszahlungsbetrag nach erreichtem Rang



**Abbildung 5.3:** Setup der xAffect-Software für die Studiendurchführung: (a) Setup mit Varioport-e, (b) Setup mit Labormesssystem

### 5.1.2 Studienablauf

Insgesamt nahmen 144 Personen an der Studie teil. Von diesen waren 113 männlich und 31 weiblich. Das Durchschnittsalter lag bei 22,86 Jahren bei einer Standardabweichung von 3,32 Jahren. Die Versuchspersonen wurden zu gleichen Teilen zufällig in zwei verschiedene Gruppen eingeteilt: High arousal (HA) und low arousal (LA).

Die Studie wurde insgesamt in 24 Sessions durchgeführt. An jeder Session nahmen jeweils 6 Versuchspersonen gleichzeitig teil. Die Sessions setzten sich zusammen aus 12 HA Sessions sowie 12 LA Sessions.

Nach Eintreffen der Teilnehmer wurden diese zunächst gebeten, anhand des SAM-Fragebogens ihre emotionale Befindlichkeit zu beurteilen. Im Anschluss daran wurden die Geräte zur Messung der physiologischen Parameter angebracht. Nachdem die Anweisungen für die erste Phase des Experiments vorgelesen wurden, folgte eine Ruhephase, in der die Teilnehmer gebeten wurden, sich möglichst ruhig zu verhalten, Bewegungen

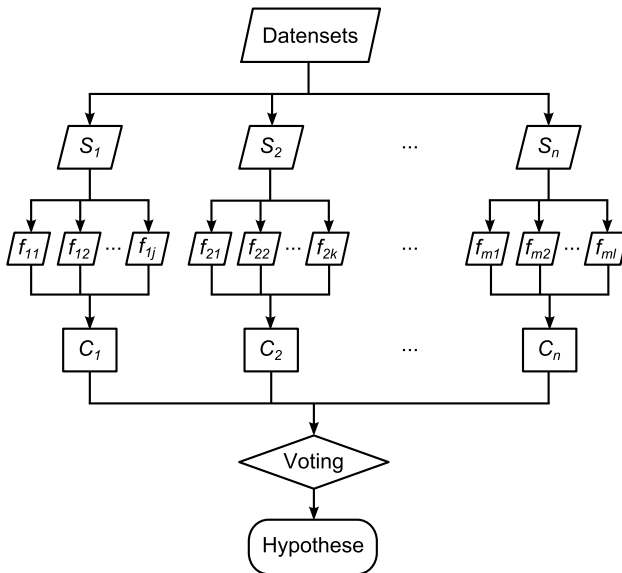
zu vermeiden und sich zu entspannen. Diese Phase diente für die spätere Datenanalyse als Baseline-Phase. Nach Abschluss der Ruhephase wurde den Teilnehmern erneut der SAM-Fragebogen vorgelegt. Es folgte eine Aufgabe zur Selbstwahrnehmung, bei der die Teilnehmer gebeten wurden, ihre Herzschläge in einem vorgegebenen Zeitintervall zu zählen. Darauf folgte eine Auktion verschiedener Gegenstände, die jedoch – wie auch die Aufgabe zur Selbstwahrnehmung – nicht Bestandteil dieser Auswertung ist. Nach der Auktion folgte der für diese Studie relevante Teil, bei welchem je nach Versuchsbedingung mittels des in Abschnitt 5.1.1.2 vorgestellten Spiels hohe oder niedrige emotionale Erregung induziert wurde. Im Anschluss an das Spiel folgte eine weitere Serie von Auktionen. Abschließend wurden die Teilnehmer erneut gebeten, den SAM-Fragebogen auszufüllen. Abgesehen von der Phase des Spiels selbst, war der Spielablauf für beide Versuchspersonengruppen gleich.

## 5.2 Datenanalyse

Nachfolgender Abschnitt dient der Analyse der in vorstehend beschriebener Studie gesammelten Daten. Im Rahmen der Vorbetrachtungen wird zunächst untersucht, inwiefern die Studie tatsächlich emotionale Erregung bei den Versuchspersonen induziert hat. Des Weiteren erfolgt eine Betrachtung der möglichen Fenstergrößen für die Online-HRV-Analyse.

In den darauf folgenden Abschnitten erfolgt die Analyse der aufgezeichneten Daten. Für alle Versuchspersonen werden für die zehnmütige Spielphase die zu untersuchenden Features in nicht überlappenden 15 Sekunden-Fenstern extrahiert. Somit stehen für die Analyse je Versuchsperson 40 Features zur Verfügung. Die Klassifikation wird personenunabhängig durchgeführt. Test und Training erfolgt nach dem leave-one-out-Verfahren, da dieses Verfahren auch im Fall einer geringen Anzahl von Datensets ein stabiles Ergebnis zulässt. Hierbei werden jeweils die Daten von  $n - 1$  Versuchspersonen zum Training und der  $n$ . Datensatz zum Testen des Klassifikators verwendet. Dies wird so lange wiederholt, bis die Daten jeder Versuchsperson einmal zum Testen verwendet wurden. Im Rahmen der Analyse werden folgende Klassifikationsverfahren verglichen: LDA, QDA, kNN, Entscheidungsbäume sowie SVM. Eine Beschreibung dieser Verfahren findet sich in Abschnitt 2.3.2. Zudem erfolgt für jedes Signal eine

Feature-Reduktion mittels Sequential Forward Selection (SFS). Für die kNN-Klassifikation wurde der Parameter  $k$  einheitlich auf 13 festgesetzt, da dieser in den meisten Analysen die besten Klassifikationsergebnisse erzielte. Da in der vorliegenden Studie ausbalancierte Daten vorliegen, wird zum Vergleich der Klassifikationsergebnisse der in Abschnitt 2.3.2.5 vorgestellte WAR (*weighted average recall*) verwendet, welcher nachfolgend der Einfachheit halber als Genauigkeit bezeichnet wird.



**Abbildung 5.4:** Klassifikationskonzept: für jedes Signal  $S$  wird ein separater Klassifikator  $C$  verwendet, welcher auf den aus dem Signal extrahierten Features  $f$  basiert. Die Verknüpfung der Klassifikatoren erfolgt über ein Voting-Verfahren

Für die Klassifikation der Daten werden zunächst alle aufgezeichneten physiologischen Signale getrennt voneinander betrachtet und hierauf entsprechende Klassifikatoren aufgebaut. Hintergrund dafür ist, dass aufgrund technischer Probleme oftmals nicht alle Signale aufgezeichnet werden konnten. Somit stehen nur wenige vollständige Datensätze zur Verfügung. Der Klassifikator wird daher kaskadierend aufgebaut: Die Klassifikation

wird zunächst für jeden Kanal separat durchgeführt. Die Ergebnisse der einzelnen Klassifikatoren werden mittels majority voting je nach Verfügbarkeit der einzelnen Kanäle verknüpft. Zur Veranschaulichung ist das Konzept der Klassifikation in Abbildung 5.4 dargestellt. Dieser Aufbau hat den Vorteil, dass alle vorhandenen Datensätze zur Klassifikation verwendet werden können, unabhängig davon, ob alle Signale aufgezeichnet wurden. Diese Vorgehensweise hat ebenfalls eine hohe Relevanz für die Online-Datenanalyse unter Alltagsbedingungen, da es hier oftmals vorkommt, dass einzelne Signale – beispielsweise durch abgefallene Elektroden – gestört oder nicht vorhanden sind.

### 5.2.1 Vorbereitung der Daten

Wie einführend dargestellt, wurden im Rahmen dieser Studie folgende Signale aufgezeichnet: EKG, PPG, EDA und Maustastendruck. Nachfolgend ist dargestellt, wie bei der Feature-Extraktion aus den aufgezeichneten Daten vorgegangen wurde. Verwendete bestehende Methoden zur Feature-Extraktion sind in Abschnitt 2.3.1 dargestellt. Weitere Verfahren, welche speziell für die hier durchgeführten Analysen entwickelt wurden, sind nachfolgend dargestellt.

#### 5.2.1.1 Kardiovaskuläre Daten

Auf dem EKG-Signal wird wie in Abschnitt 2.3.1 dargestellt zunächst eine Detektion der R-Zacken vorgenommen. Bevor hieraus alle weiteren Features extrahiert werden, wird für die Offline-Analyse zusätzlich zu der automatischen Bereinigung fehlerhaft detektierter R-Zacken eine visuelle Inspektion der Daten vorgenommen, da fehlerhafte R-Zacken – insbesondere bei den frequenzbasierten HRV-Parametern – zu einer starken Verfälschung der Analyse-Ergebnisse führen können.

Vor der Feature-Extraktion aus dem PPG-Signal sollen zunächst hochfrequente Störungen aus dem Signal entfernt werden. Es wird daher eine Glättung des Signals mit einem Hann-Fenster der Größe von 16 Samples durchgeführt, was einer Dauer von 128 ms entspricht. Die Fenstergröße für die Glättung wurde möglichst klein gewählt, um die durch die Glättung bedingte zeitliche Verschiebung des Signals möglichst gering zu halten, damit die Merkmalszeitpunkte des Signals möglichst genau

bestimmt werden können. Zusätzlich erfolgt eine Filterung des Signals mit einem Butterworth-Hochpassfilter 2. Ordnung mit einer Grenzfrequenz von 0,15 Hz. Mit Hilfe des in Abschnitt 2.3.1.1 dargestellten Verfahrens wurden auf dem PPG-Signal Minimum, Maximum sowie der virtuelle Fußpunkt detektiert. Basierend auf Minimum und Maximum wird anschließend die Amplitudenhöhe ermittelt. Zusätzlich wird in jedem Zeitfenster die Standardabweichung der Amplitudenhöhe berechnet. Die PTT wird als Differenz zwischen der zuvor aufgetretenen R-Zacke und dem virtuellen Fußpunkt bestimmt.

Im Vergleich zum EKG-Signal treten im PPG-Signal verhältnismäßig häufig Störungen auf, da diese bei Messung am Finger bereits durch leichte Bewegungen des Fingers verursacht werden. Aus diesem Grund werden bei der Detektion der Pulswellencharakteristika zusätzlich Kontextinformationen aus dem EKG-Signal herangezogen: Eine Pulswelle wird nur dann als valide gewertet, wenn sie sich innerhalb der von Boll (2007) definierten Zeitgrenzen befindet (siehe hierzu Abschnitt 2.3.1.1).

Die kardiovaskulären Features, welche im Rahmen der nachfolgenden Analysen betrachtet werden, sind in Tabelle 5.2 dargestellt.

Signal	Analysebereich	Feature	Einheit
EKG	Zeit	Herzrate	bpm
		SDNN	ms
		RMSSD	ms
		SD1	ms
		SD2	ms
		SD1/SD2-Ratio	–
		pNN12	%
		pNN20	%
		pNN50	%
	Frequenz	LF	ms <sup>2</sup>
		HF	ms <sup>2</sup>
		LF/HF-Ratio	–
	PPG	Zeit	Amplitudenhöhe
Std Amplitudenhöhe			V
PPG&EKG	Zeit	PTT	ms

**Tabelle 5.2:** Kardiovaskuläre Features für die Datenanalyse



### 5.2.1.2 Elektrodermale Daten

Zur Analyse der elektrodermalen Aktivität wird im ersten Schritt ein Downsampling des Rohsignals auf 32 Hz vorgenommen. Diese Genauigkeit kann für die genaue Bestimmung der EDA-basierten Features als ausreichend angesehen werden und weitere Berechnungen werden hierdurch beschleunigt.

Zusätzlich wird vor der Detektion der Reaktionen zunächst eine Tiefpassfilterung des Signals mittels eines Butterworth-Filters 2. Ordnung und einer Grenzfrequenz von 1 Hz vorgenommen, um hochfrequente Störungen zu eliminieren.

In Anlehnung an die in der Literatur häufig verwendeten Features, werden die in Tabelle 5.3 aufgeführten Features mit in die Analysen einbezogen. Die eventbasierten Features beziehen sich hierbei auf einzelne elektrodermale Reaktionen während die zeitbasierten Features auf Signalbereichen berechnet werden.

Analysebereich	Feature	Einheit
Events	Amplitudenhöhe	$\mu\text{S}$
	Anstiegszeit	ms
	Halbwertszeit	ms
	Signalenergie	$\text{ms}^2$
Zeit	Anzahl SCRs	–
	Anzahl Nulldurchgänge	–
	Amplitude Nulldurchgänge	$\mu\text{S}$
	SCL	$\mu\text{S}$
	positive Steigung	$\mu\text{S}$
	Aktivität	–
	Mobilität	–
Komplexität	–	

**Tabelle 5.3:** EDA-Features für die Datenanalyse

### 5.2.1.3 Kraftdaten

Von den Kraftsensoren, die in der in Abschnitt 4.2.2 beschriebenen kraftsensitiven Maus integriert sind, wird ein kontinuierliches Kraftsignal übertragen. Zur weiteren Analyse der auf die Maustaste ausgeübten Kraft müssen daher zunächst die Mausklicks detektiert werden.

Im ersten Schritt erfolgt eine Normalisierung der Daten. Hierfür wird der Modalwert des gesamten Signals von dem Signal abgezogen. Dies ist erforderlich, da durch die Maustaste selbst eine konstante Kraft auf den Drucksensor ausgeübt wird. Anschließend erfolgt eine Filterung des Signals mit einem Butterworth-Hochpassfilter 2. Ordnung mit einer Grenzfrequenz von 2 Hz, der auf das normalisierte Signal angewendet wird, um niederfrequente Veränderungen herauszufiltern. Im letzten Schritt werden die Maxima des Signals mit Hilfe der 1. Ableitung detektiert.

Da es vorkommen kann, dass Kraft auf die Maustaste ausgeübt wird, jedoch kein Klick ausgelöst wird, werden nur Maxima als valide Mausklicks gewertet, die mindestens 10 % des Maximums des Gesamtsignals betragen.

Basierend auf den detektierten Klickzeitpunkten wird die Klickstärke als das Maximum des jeweiligen Mausklicks ermittelt. Zusätzlich werden für jeden Mausklick Schiefe und Wölbung des dazugehörigen Kraftsignals ermittelt, da diese beiden Parameter weiteren Aufschluss über das Klickverhalten einer Person geben können. Die Schiefe des Signals gibt an, wie symmetrisch ein Mausklick ist, d. h. ob die Maustaste beispielsweise schneller gedrückt als losgelassen wurde. Die Wölbung gibt Aufschluss über die Steilheit des Mausklicks. Kurze, schnelle Mausklicks sollten in der Regel steiler ausfallen, als langsame Mausklicks.

Tabelle 5.4 zeigt eine Zusammenfassung der aus dem Kraftsignal extrahierten Features.

Analysebereich	Feature	Einheit
Events	Kraft auf Maustaste	N
	Schiefe des Mausklicks	–
	Wölbung des Mausklicks	–

**Tabelle 5.4:** Maus-Features für die Datenanalyse

Für die Fensterung wird jeweils der Mittelwert aller in einem Zeitfenster aufgetretenen Klicks gebildet.

Die Daten des seitlich an der Maus angebrachten Kraftsensors wurden nicht weiter in die Analysen mit einbezogen, da aufgrund der Platzierung keine aussagekräftigen Daten aufgezeichnet wurden.

## 5.2.2 Vorabbetrachtungen

Vor der eigentlichen Datenanalyse werden zunächst zwei Vorabanalysen durchgeführt. Die erste Analyse betrifft die Frage, inwiefern durch das Spiel tatsächlich Erregung induziert wurde. Ferner wird für die weiteren Analysen untersucht, bis zu welcher Fenstergröße HRV-Parameter sinnvoll zur Erkennung emotionaler Erregung verwendet werden können.

### 5.2.2.1 Validität der Erregungsinduktion

Vor der Analyse der physiologischen Daten muss zunächst geprüft werden, ob durch die unterschiedlichen Spielszenarien tatsächlich emotionale Erregung induziert werden konnte. Hierfür werden die SAM-Ratings verwendet, die sowohl nach der Baseline-Phase vor dem Spiel (I) als auch nach Beendigung des Spieles (II) erhoben wurden.

Tabelle 5.5 veranschaulicht, dass sich die Mittelwerte des SAM-Ratings hinsichtlich der emotionalen Erregung zwischen den Gruppen nach dem Spiel erheblich unterscheiden während vor dem Spiel keine Mittelwertsdifferenzen vorliegen. Dies bestätigt auch der Wilcoxon-Rangsummentest: Während für das SAM-Rating nach der Baseline-Phase vor dem Spiel keine signifikanten Unterschiede zu beobachten sind ( $W = 5138,5$ ;  $Z = -0,342$ ;  $p = 0,733$ ), zeigen sich nach dem Spiel hochsignifikante Unterschiede für beide Bedingungen ( $W = 3952,0$ ;  $Z = -5,126$ ;  $p < 0,001$ ).

Dies bedeutet, dass die Unterschiede zwischen den Erregungsniveaus vor dem Spiel für die HA-Bedingung signifikant stärker sind als für die LA-Bedingung. Im Gegensatz hierzu konnten für die Valenz-Ratings weder vor ( $W = 5214,5$ ;  $Z = -0,023$ ;  $p < 0,982$ ) noch nach dem Spiel ( $W = 4989,0$ ;  $Z = -0,953$ ;  $p < 0,341$ ) signifikante Unterschiede festgestellt werden.

	Valenz		Erregung	
	HA	LA	HA	LA
(I)	6,29 (1,32)	6,31 (1,26)	2,10 (1,06)	2,21 (1,51)
(II)	6,44 (1,35)	6,26 (1,39)	4,61 (1,95)	2,96 (1,45)

**Tabelle 5.5:** Mittelwerte und Standardabweichung für Valenz- und Erregungsratings vor und nach dem Spiel

Die Analyse der SAM-Ratings zeigt, dass durch die Intervention mittels

des Spiels emotionale Erregung bei den Versuchspersonen induziert wurde, während hinsichtlich der Valenz keine Veränderungen zu beobachten sind.

### 5.2.2.2 Evaluation der Fenstergröße für die HRV-Analyse

Für die Online-Klassifikation von Erregung ist es wichtig, dass die Datenanalyse ohne großen Zeitversatz erfolgt. Aus diesem Grund wurden bei EKG-basierten Analysen in der Vergangenheit in erster Linie Features basierend auf der Herzrate betrachtet. HRV-Features wurden zumeist vernachlässigt, da die Fenstergröße für die Analysen gemäß der Empfehlungen der Taskforce (Malik, 1996) selbst für Kurzzeit-Analysen bei mindestens 5 Minuten liegen sollte. Erste Studien mit sogenannten Ultra-Kurzzeit Zeitfenstern unter einer Minute haben jedoch gezeigt, dass einige HRV-Parameter auch für kleinere Zeitfenster zuverlässige Ergebnisse liefern. So konnten Schroeder et al. (2004) hohe Korrelationen zwischen 2 und 6 Minuten-Fenstern für die mittlere HR, den mittleren RR-Abstand, SDNN sowie RMSSD feststellen, wobei die Korrelationen zwischen 10 Sekunden und 6 Minuten-Fenstern für SDNN und RMSSD etwas geringer ausfielen. Für frequenzbasierte Parameter konnten lediglich geringere Korrelationen festgestellt werden. Nussinovitch et al. (2011) fanden hohe Korrelationen zwischen 10 Sekunden- und 5 Minuten-Fenstern für RMSSD und den mittleren RR-Abstand, während für den maximalen und minimalen RR-Abstand, SDNN, HRV triangular index, NN50, pNN50 sowie die frequenzbasierten HRV-Parameter keine signifikanten Korrelationen gefunden wurden. Die Eignung dieser Merkmale hinsichtlich der Erkennung unterschiedlicher Erregungsniveaus wurde bisher nicht betrachtet.

Ziel der Evaluation der Fenstergröße ist die Bestimmung einer minimalen Fenstergröße für die einzelnen HRV-Features gewählt werden kann, welche noch eine zuverlässige Klassifikation des Erregungsniveaus ermöglicht. In Anlehnung an die vorhandene Literatur wird bei der HRV-Analyse im Folgenden bei einem Zeitfenster von 5 Minuten der Begriff der Kurzzeit-Analyse verwendet und bei kleineren Zeitfenstern von Ultra-Kurzzeit-Analyse gesprochen.

Für die HRV-Features im Frequenzbereich ergeben sich für die zu evaluierenden Fenstergrößen einige Einschränkungen. Bei einer Fenstergröße von  $T$  Sekunden liegt die niedrigste auflösbare Frequenz bei  $1/T$  Hz. Für die HRV-Frequenzbänder ergeben sich somit folgende Untergrenzen für

die Zeitfenster:

Frequenzband	Untergrenze [Hz]	Zeitfenster [s]
VLF	0,003	333,33
LF	0,04	25,00
HF	0,15	6,67

**Tabelle 5.6:** Minimale Zeitfenster für die Analyse von HRV-Frequenzbändern in Abhängigkeit von der Untergrenze des Frequenzbands

Aufgrund der in Tabelle 5.6 dargestellten Anforderungen an die minimale Fenstergröße, wird VLF bei den weiteren Analysen nicht betrachtet. Bei LF sowie dem LF/HF-Ratio wird als Untergrenze für die Analysen 30 Sekunden festgelegt. Für alle anderen HRV-Features wird als Untergrenze für die Analyse ein Zeitfenster von 15 Sekunden betrachtet. Die jeweiligen HRV-Parameter für die einzelnen Fenstergrößen werden aus der Mitte des Spiels extrahiert, welches eine Gesamtdauer von 10 Minuten hat. Hierdurch soll die Vergleichbarkeit der HRV-Parameter gewährleistet werden, da die Parameter auf diese Weise nicht durch eine eventuelle Anspannung zu Beginn und Ende des Spiels verfälscht werden.

#### Zusammenhang zwischen ultrakurzen und kurzen Zeitfenstern

Für die nachfolgenden Analysen wurden die Fenstergrößen 15, 30, 60 sowie 300 Sekunden untersucht. Die Mittelwerte und Standardabweichungen der einzelnen Features sind der Tabelle in Anhang A zu entnehmen. Es zeigt sich, dass die Abweichungen für die unterschiedlichen Fenstergrößen bei den einzelnen Features unterschiedlich groß ausfallen. Eine ANOVA zeigt, dass sich die Werte der Features für die unterschiedlichen Fenstergrößen bei RMSSD, pNNx, SD1 sowie dem LF/HF-Ratio nicht signifikant unterscheiden. Jedoch sind für die Features SDNN, SD2, SD1/SD2-Ratio sowie LF und HF deutliche Unterschiede zwischen den Werten ersichtlich. Tabelle 5.7 zeigt die Ergebniss der ANOVA für die einzelnen Features sowie den dazugehörigen Standardfehler. Des Weiteren zeigt Tukey's HSD<sub>95%</sub> auf einem Alpha-Niveau von 5% die Schwellwerte an, ab welchen der Unterschied zwischen zwei Mittelwerten signifikant wäre. Interessant ist, dass die Unterschiede zwischen den Feature-Mittelwerten für die LA-Bedingung etwas geringer ausfallen als für die HA-Bedingung. Eine mögliche Erklä-

rung für die leicht geringeren Unterschiede kann darin gesehen werden, dass bei den Signalen unter niedrigerer emotionaler Erregung insgesamt eine geringere Variabilität zu erwarten ist.

	LA			HA		
	$F$	SE	HSD <sub>95%</sub>	$F$	SE	HSD <sub>95%</sub>
SDNN	3,207*	3,905	10,108	7,192***	3,013	7,796
RMSSD	0,175	3,511	9,086	0,271	2,883	7,461
pNN12	0,029	0,037	0,097	0,077	0,051	0,131
pNN20	0,049	0,041	0,101	0,068	0,045	0,116
pNN50	0,083	0,029	0,074	0,006	0,021	0,054
SD1	0,181	2,478	6,414	0,052	2,069	5,353
SD2	5,584*	5,102	13,206	10,901***	3,866	10,005
SD1/SD2	13,697***	0,033	0,087	10,504***	0,043	0,110
LF	2,114	0,023	0,055	13,514***	0,027	0,064
HF	8,595***	0,016	0,038	10,941***	0,016	0,038
LF/HF	0,015	0,503	1,189	0,264	0,751	1,776

\*  $p < 0,05$ ; \*\*  $p < 0,01$ ; \*\*\*  $p < 0,001$

**Tabelle 5.7:** Ergebnisse der ANOVA für die jeweiligen experimentellen Bedingungen

### Trennschärfe der HRV-Parameter

Wie die Ergebnisse bezüglich der Zusammenhänge zwischen ultrakurzen und kurzen Fenstergrößen zeigen, gibt es große Unterschiede zwischen den einzelnen HRV-Parametern, in welchem Maß ultrakurze Zeitfenster ähnliche Informationen abbilden wie kurze Zeitfenster. Im nächsten Schritt wird daher untersucht, ob die jeweiligen Parameter zur Diskriminierung zwischen hoher und niedriger emotionaler Erregung geeignet sind. Die Mittelwerte für die einzelnen HRV-Features in der Tabelle in Anhang A deuten darauf hin, dass die frequenzbasierten Features unterschiedlich gut zur Differenzierung zwischen hoher und niedriger Erregung geeignet sind, da bei einigen Features die Mittelwertsdifferenzen für hohe und niedrige Erregung sehr gering ausfallen.

Zur weiteren Quantifizierung der Eignung der HRV-Features zur Unterscheidung zwischen zwei Erregungszuständen wird daher Fisher's Discriminant Ratio (FDR) für die einzelnen Features sowohl für kurze als auch für ultrakurze Zeitfenster herangezogen. Das FDR beschreibt das Verhältnis

der Varianz zwischen den Klassen zu der Varianz innerhalb der Klassen und kann wie folgt berechnet werden:

$$FDR = \frac{(\mu_{HA} - \mu_{LA})^2}{\sigma_{HA}^2 + \sigma_{LA}^2}$$

wobei  $\mu_{HA}$  und  $\mu_{LA}$  die Mittelwerte und  $\sigma_{HA}^2$  und  $\sigma_{LA}^2$  die Standardabweichungen für die beiden Klassen LA und HA angeben. Je höher der Wert, den das FDR annimmt, desto eher ist ein entsprechendes Feature dazu geeignet, zwischen den zwei Klassen zu differenzieren.

Da es große Unterschiede hinsichtlich der HRV-Parameter verschiedener Personen gibt, wird vor der Berechnung des FDR eine personenabhängige Normalisierung der Daten anhand der Daten aus der Baseline-Periode vorgenommen. Die normalisierten Features  $f_{norm}$  werden wie folgt berechnet:

$$f_{norm} = \frac{f_{game} - f_{base}}{f_{base}}$$

Für einige Versuchspersonen traten für pNN20 und pNN50 Nullwerte während der Baseline-Periode auf. Diese Features wurden von der weiteren Analyse ausgeschlossen. Es ergeben sich daher folgende Werte für die Anzahl der Features für pNN20 und pNN50:  $N_{HA} = 58$  sowie für pNN50  $N_{LA} = 53$  während für alle anderen Feature weiterhin gilt  $N_{HA} = 59$  und  $N_{LA} = 58$ .

	$FDR_{15}$	$FDR_{30}$	$FDR_{60}$	$FDR_{300}$
SDNN	0.337	0.443	0.404	0.318
RMSSD	0.376	0.376	0.419	0.321
pNN12	0.290	0.318	0.313	0.377
pNN20	0.098	0.101	0.061	0.097
pNN50	0.025	0.013	0.014	0.017
SD1	0.373	0.424	0.405	0.323
SD2	0.298	0.402	0.366	0.303
SD1/SD2	0.000	0.002	0.011	0.040
HF	-	0.020	0.004	0.007
LF	-	0.008	0.003	0.163
LF/HF	-	0.000	0.000	0.007

**Tabelle 5.8:** Fisher’s discriminant Ratio für alle Features für die unterschiedlichen Fensterlängen

Die in Tabelle 5.8 dargestellten Ergebnisse zeigen, dass es große Unterschiede bezüglich der Diskriminierungsfähigkeit zwischen den einzelnen HRV-Parametern gibt. Von den betrachteten HRV-Features sind SDNN, RMSSD sowie SD1 am besten geeignet, um zwischen den beiden Klassen zu differenzieren. Aufgrund geringer Werte des FDR bei den frequenzbasierten Features kann davon ausgegangen werden, dass diese für die Daten aus der durchgeführten Studie nur wenig zur personenunabhängigen Unterscheidung zwischen den beiden Klassen geeignet sind.

### Zusammenfassung

Die vorstehenden Analysen zeigen, dass auch die HRV zur Klassifikation emotionaler Erregung in Zeitfenstern, welche kleiner als die üblicherweise verwendeten 5 Minuten sind, geeignet ist. Jedoch schwankt die Eignung je nach betrachtetem Feature stark. Insbesondere pNN12, pNN20, RMSSD und SD1 stellen sich als sehr geeignet für die Klassifikation von emotionaler Erregung in ultrakurzen Zeitfenstern von 15 Sekunden dar, während die frequenzbasierten Features für die Ultrakurzzeit-Analyse ungeeignet erscheinen.

## 5.2.3 Analyse der kardiovaskulären Aktivität

Die kardiovaskuläre Aktivität wurde im Rahmen dieser Studie sowohl mit EKG- als auch mit PPG-Sensoren erfasst. Dies ermöglicht die Analyse vielseitiger kardiovaskulärer Parameter.

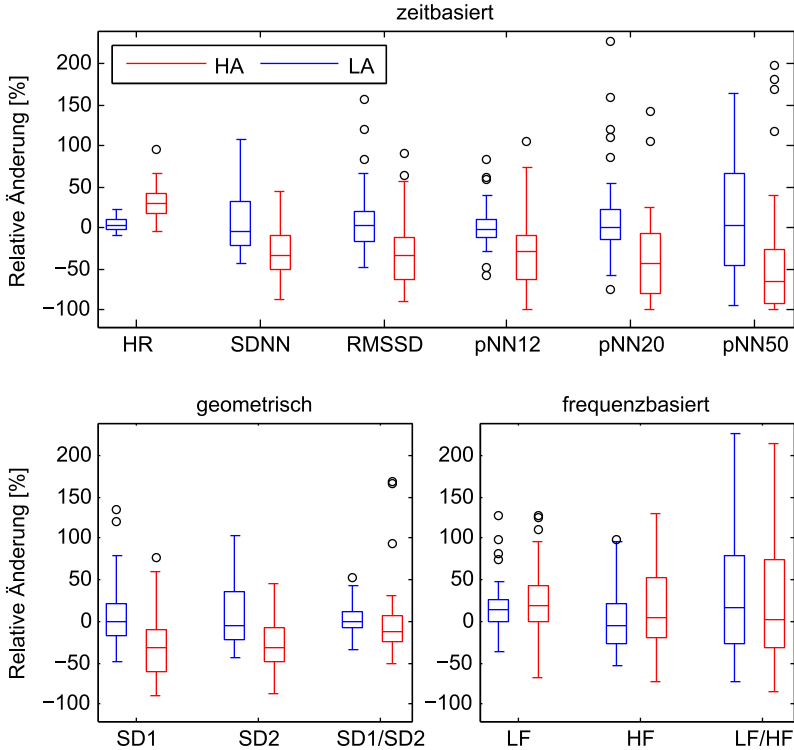
Da die Signalqualität der EKG-Daten im Vergleich zu den PPG-Daten weitaus besser ausfällt, sowie der aufgrund der genaueren Möglichkeiten zur Bestimmung des exakten Zeitpunkts eines Herzschlages, werden für alle Features, welche auf den NN-Abständen basieren, ausgehend von dem EKG-Signal berechnet.

### 5.2.3.1 EKG-basierte Features

Nachfolgend werden die Mittelwerte der einzelnen EKG-basierten Features aufgeteilt nach experimentellen Bedingungen untersucht. Für die Untersuchungen standen für die HA-Bedingung 59 und für die LA-Bedingung 58 vollständige Datensätze zur Verfügung. Um Zwischensubjektunterschiede auszugleichen, wird eine Normalisierung der Features vorgenommen: Für



jeden Datensatz wird die relative Abweichung des Mittelwerts über die 15-Sekundenfenster (bzw. 30-Sekundenfenster für frequenzbasierte Features) während des Spiels von dem Mittelwert der Baseline dargestellt.



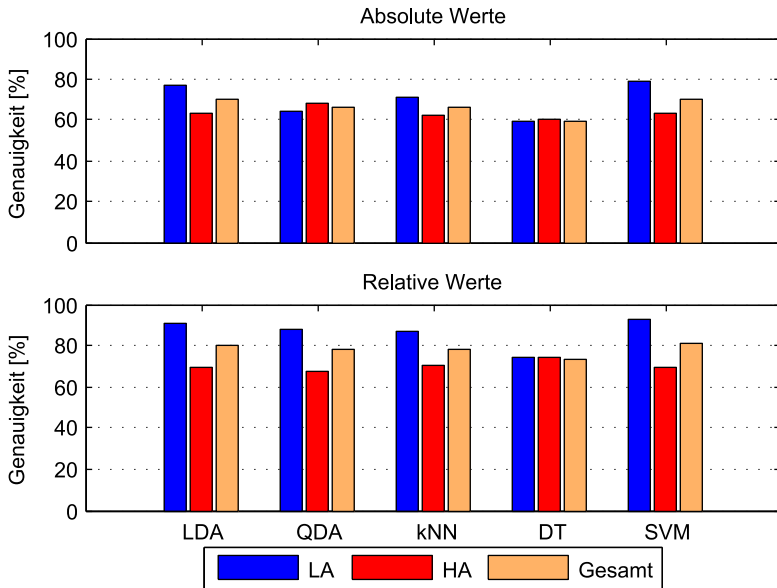
**Abbildung 5.5:** Relative Veränderung der EKG-basierten Features im Vergleich zur Baseline

Die Boxplots in Abbildung 5.5 zeigen die Verteilungen der aus dem EKG-Signal berechneten Features untergliedert nach Features im Zeitbereich, geometrischen Features und frequenzbasierten Features. Zwar sind die frequenzbasierten Features wie in Abschnitt 5.2.2.2 erläutert nur bedingt zur Online-Klassifikation von Daten geeignet, jedoch sind diese der Vollständigkeit halber mit aufgeführt. Für die weiteren Analysen werden sie jedoch aufgrund ihrer schlechten diskriminatorischen Eigenschaften nicht

weiter betrachtet.

Die Abbildung bestätigt, was sich bereits in den Vorabanalysen zeigte: Die größten Unterschiede zwischen den einzelnen Erregungsstufen sind für die zeitbasierten Features erkennbar. Auch die geometrischen Features spiegeln das Erregungsniveau der Versuchspersonen wider. Jedoch sind anhand der frequenzbasierten Features keine Unterschiede zwischen den einzelnen Klassen erkennbar.

Basierend auf den zeitbasierten und den geometrischen Features erfolgt im nächsten Schritt die Klassifikation der Daten. Abbildung 5.6 zeigt die Ergebnisse der Klassifikation mit allen Features. Zusätzlich zu den relativen Features ist zum Vergleich auch das Ergebnis der Klassifikation mit den absoluten Features dargestellt. Wie erwartet, fällt die Klassifi-



**Abbildung 5.6:** Klassifikationsgenauigkeiten verschiedener Klassifikatoren für emotionale Erregung auf EKG-basierten Features

kationsgenauigkeit für die absoluten Features deutlich geringer aus, als für die relativen Features. Die besten Ergebnisse für die Klassifikation

können mittels LDA und SVMs erzielt werden. Für relative Features wird hier eine Gesamtgenauigkeit von 80,45 % für die LDA und 80,91 % für SVMs erreicht. Der große Vorteil der LDA gegenüber SVMs liegt in der weitaus geringeren Komplexität des Verfahrens und den somit für die Online-Klassifikation weitaus günstigeren Laufzeiteigenschaften.

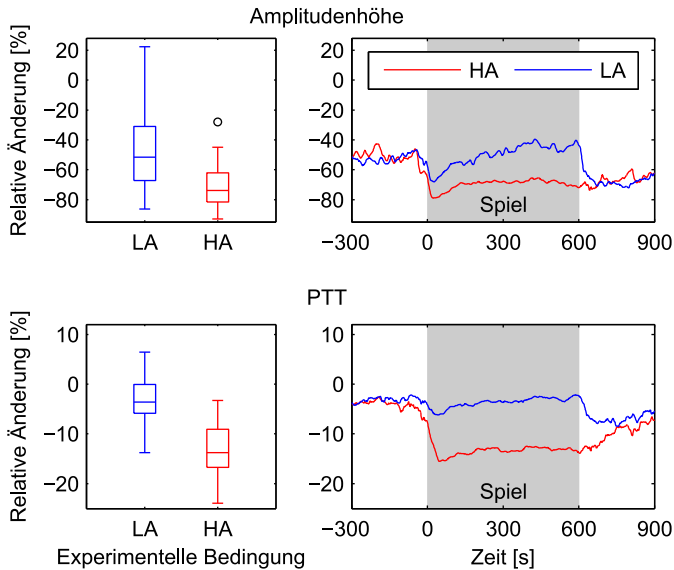
Zusätzlich wird eine Feature-Reduktion mittels SFS vorgenommen. Diese liefert als optimales Ergebnis eine Feature-Kombination aus Herzrate, RMSSD und SD1. Mittels dieser Feature-Kombination kann die Erkennungsrate bei der LDA auf 80,69 % verbessert werden. Für die SVM-Klassifikation bleibt die Klassifikationsgenauigkeit auch nach der Feature-Reduktion bei 80,91 % .

### 5.2.3.2 PPG-Amplitude und PTT

Für die Analyse der PPG-Daten wurden nur Datensätze berücksichtigt, bei welchen während der Baseline und während des Spiels mindestens 90 % des Signals artefaktfrei waren. Aufgrund der hohen Artefaktanfälligkeit des PPG-Sensors bei der Messung am Finger sowie technischer Probleme bei drei der Messgeräte, war nur ein weitaus geringerer Teil der Daten verwertbar, als bei den EKG-Daten. Nachfolgend liegt die Stichprobengröße daher bei  $N_{HA} = 23$  und  $N_{LA} = 31$ .

In Abbildung 5.7 ist die relative Veränderung von PTT und PPG-Amplitude im Vergleich zur Baseline für hohe und niedrige Erregung im Mittel über alle Datensätze dargestellt. Es zeigt sich, dass insbesondere bei PTT eine sehr klare Abgrenzung zwischen hoher und niedriger Erregung vorhanden ist. Während des Spiels ist die PTT für die HA-Bedingung durchgängig mindestens 10 % niedriger als während der Baseline, wohingegen die Veränderung der PTT für die LA-Bedingung über den kompletten Spielverlauf gering bleibt. Zwar fällt die relative Veränderung der Amplitudenhöhe weitaus größer aus, als die der PTT, jedoch befindet sich die Amplitudenhöhe bereits zu Beginn des Spiels deutlich unter dem Baseline-Niveau. Die Ursache hierfür kann darin gesehen werden, dass der Sensor die Lichtintensität im Verlauf der Messung immer wieder adaptiert. Umso erstaunlicher ist es, dass sich trotzdem Unterschiede zwischen beiden Versuchsbedingungen erkennen lassen.

Für die Klassifikation auf der Kombination aus PPG- und EKG-Features reichen die Klassifikationsgenauigkeiten von 74,07 % für Entscheidungs-



**Abbildung 5.7:** Relative Veränderung von PPG-Amplitude (oben) und PTT (unten) im Vergleich zur Baseline gemittelt nach experimenteller Bedingung. Links: Boxplot der Mittelwerte während des Spiels, Rechts: Relativer Signalverlauf während des Spiels

bäume bis hin zu 80,66 % für die LDA. Erneut schneiden SVMs mit einer Genauigkeit von 79,66 % nur wenig schlechter ab. kNN hingegen erzielen nur eine Genauigkeit von 75,57 %. Aus der SFS ergibt sich eine optimale Featurekombination bestehend aus PTT und Amplitudenhöhe. Für die LDA steigt die Genauigkeit hiermit auf 81,53 %, für SVMs auf 80,12 %.

### 5.2.4 Analyse der EDA

Wie bereits bei der EKG-Analyse wird auch bei der EDA-Analyse eine Fensterlänge von 15 Sekunden zugrunde gelegt. Bei den eventbasierten Features wird jeweils der Mittelwert aller in dem analysierten Zeitfenster aufgetretenen Events berechnet. Eine detaillierte Beschreibung der verwendeten Features sowie deren Berechnung findet sich in Abschnitt

2.3.1.2. Aufgrund technischer Probleme stehen für die Analysen insgesamt 44 Datensätze zur Verfügung, welche sich gleichmäßig auf beide Versuchsbedingungen aufteilen.

In Tabelle 5.9 sind Mittelwerte und Standardabweichungen der analysierten EDA-Features während des Spiels im Vergleich zur Baseline dargestellt. Es zeigt sich, dass die Mittelwerte für beide Versuchsbedingungen sehr dicht beieinander liegen, was eine Trennung zwischen den Klassen erschwert.

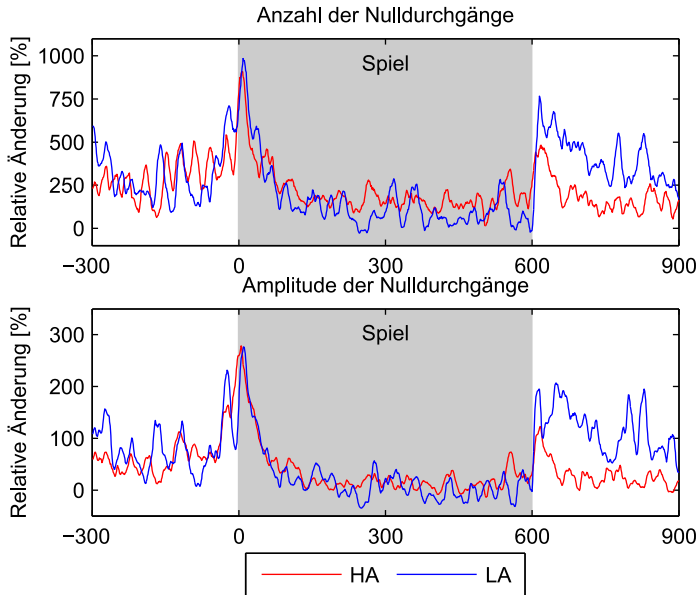
Feature	LA	HA
Amplitudenhöhe	-22,368 (12,763)	-20,725 (11,926)
Signalenergie	-20,741 (13,757)	-21,854 (13,241)
Halbwertszeit	6,362 (8,832)	-1,010 (8,069)
Anstiegszeit	-0,058 (5,032)	3,556 (4,603)
Anzahl SCR	-43,779 (68,734)	5,823 (179,020)
Anzahl Nulldurchgänge	163,375 (263,393)	215,721 (493,283)
Aktivität	223,748 (664,994)	148,010 (405,481)
Mobilität	46,837 (53,141)	69,699 (85,018)
Komplexität	-42,348 (32,225)	-51,436 (32,732)
SCL	26,065 (27,099)	44,385 (65,365)
Positive Steigung	163,358 (364,330)	255,814 (524,842)
Amplitude Nulldurchgänge	27,500 (104,629)	35,664 (83,825)

**Tabelle 5.9:** Mittlere prozentuale Veränderung der EDA-Features während des Spiels im Vergleich zur Baseline (Standardabweichung in Klammern)

Die Klassifikation wird getrennt nach zeit- und eventbasierten Features durchgeführt, da für Zeitfenster, in welchen keine SCRs auftreten, keine eventbasierten Features berechnet werden können und somit eine geringere Anzahl an Featurevektoren zur Verfügung steht.

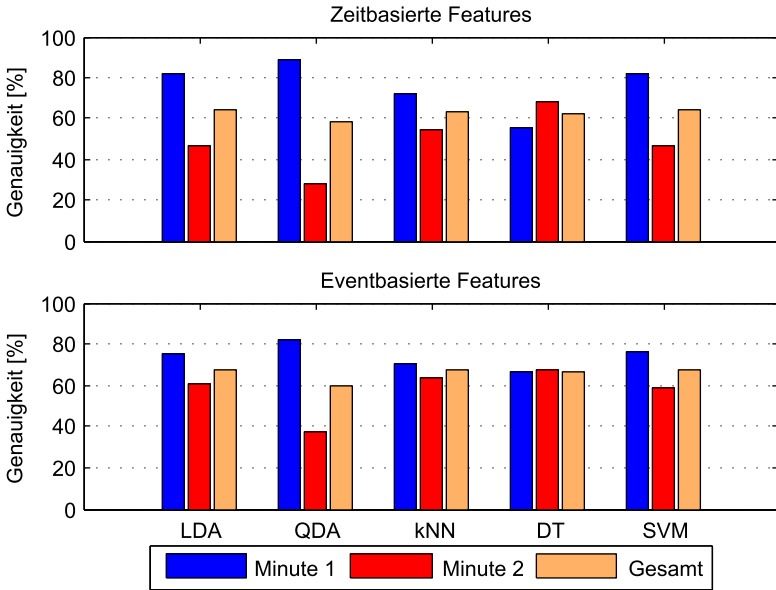
Erneut werden bei der Klassifikation LDA, QDA, kNN, Entscheidungsbäume und SVMs verglichen. Zudem erfolgt eine Optimierung der verwendeten Features mittels SFS. Für die zeitbasierten Features ergibt sich die beste Klassifikationsgenauigkeit aus der Kombination des mittleren SCL sowie der Anzahl der Nulldurchgänge. Bei den eventbasierten Features kann die beste Genauigkeit mit der mittleren Amplitudenhöhe erreicht werden. Jedoch bewegen sich die Genauigkeiten für alle Klassifikatoren sowohl bei den zeit- als auch bei den eventbasierten Features dicht um das Zufallsniveau von 50,00 %. Bei den zeitbasierten Features schwankt die Genauigkeit von 45,76 % für kNN bis 56,53 % für die LDA, bei den event-

basierten Features von 45,51 % für Entscheidungsbäume bis 53,42 % für die LDA. Auch eine Kombination beider Klassifikatoren mittels majority voting bringt keine Verbesserung mit sich.



**Abbildung 5.8:** Relative Veränderung von Amplitude und Anzahl der Nulldurchgänge während des Spiels im Vergleich zur Baseline

Ein Blick auf den Signalverlauf der einzelnen Features während des Spiels verdeutlicht, dass sich die Verläufe für hohe und niedrige Erregung während des Spiels nicht wesentlich unterscheiden. Abbildung 5.8 zeigt beispielhaft die Verläufe für die Anzahl und die Amplitude der Nulldurchgänge im Mittel für die jeweilige Versuchsbedingung. Es ist deutlich zu erkennen, dass bei beiden Signalen – unabhängig von der Versuchsbedingung – eine große Veränderung zu Beginn des Spiels stattfindet, welche jedoch schnell wieder auf das Ausgangsniveau zurückkehrt. Aus diesem Grund wird in einem nächsten Schritt untersucht, inwieweit der Spielbeginn von dem restlichen Spielverlauf mittels eines Klassifikators getrennt werden kann. Hierzu wird die erste Klasse aus den vier Samples während des Spiels über



**Abbildung 5.9:** Klassifikationsgenauigkeiten verschiedener Klassifikatoren für emotionale Erregung basierend auf dem Zeitverlauf des EDA-Signals

beide Erregungsniveaus gebildet, was der ersten Spielminute entspricht. Die zweite Klasse wird aus der zweiten Spielminute gewählt. Somit sind die Samples gleichmäßig über beide Klassen verteilt. Zusätzlich wird mittels SFS eine Feature-Selektion durchgeführt, durch welche eine weitere Verbesserung der Klassifikationsgenauigkeit erfolgt. Für die zeitbasierten Features setzt sich das optimale Feature-Set aus dem mittleren SCL, der mittleren Amplitude der Nulldurchgänge sowie der mittleren positiven Steigung zusammen. Für die eventbasierten Features ergibt sich ein zweidimensionales Featureset bestehend aus der mittleren Amplitudenhöhe sowie der mittleren Anstiegszeit. Abbildung 5.9 zeigt die Erkennungsraten der verschiedenen Klassifikatoren für die zeit- und eventbasierten Features. Wie auch schon bei der Herzrate kann die höchste Klassifikationsgenauigkeit mittels LDA erzielt werden. Für die zeitbasierten Features liegt diese

bei 64,39 %, für die eventbasierten Features bei 67,82 %. Auffällig ist die schlechte Eignung von SVMs für diesen Fall, da bei beiden Klassen fast alle Features derselben Klasse zugeordnet werden. Durch eine Kombination des event- und des zeitbasierten Klassifikators kann keine weitere Verbesserung der Klassifikationsgenauigkeit erzielt werden.

Die Datenlage zeigt, dass EDA nur bedingt zur Identifikation längerandauernder Erregungszustände geeignet ist. Die Ursache hierfür kann in den physiologischen Eigenschaften der EDA gesehen werden: Bei häufigem Auftreten von Responses kommt es schnell zu einer Sättigung, wodurch die Amplitudenhöhe stark abnimmt und Responses oftmals so gering ausfallen, dass sie nicht mehr detektiert werden können, da sie unterhalb der Detektionsgrenzen liegen und oftmals von Artefakten überlagert sind. Auch beim SCL ist schnell ein Sättigungsniveau erreicht, so dass hier kein weiterer Anstieg mehr erfolgt.

Da die Erkennung plötzlicher Zustandsänderungen auch für Lernanwendungen von Bedeutung sein kann, wird dieser Aspekt in Abschnitt 6 weiter beleuchtet. Das Erkennen einer plötzlich auftretenden Zustandsänderung könnte während des Lernens beispielsweise dazu führen, dass das Lernprogramm automatisch kurzzeitig unterbrochen wird.

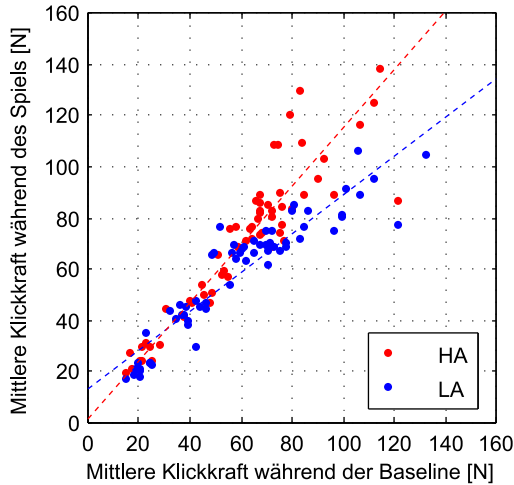
## 5.2.5 Analyse der auf die Maus ausgeübten Kraft

Neben der Erfassung der kardiovaskulären und elektrodermalen Aktivität wurde zusätzlich die Erfassung der auf die Maustaste ausgeübten Kraft mittels der in Abschnitt 4.2.2 vorgestellten Maus erfasst. Insgesamt standen für die Analyse 124 Datensätze mit Kraftdaten zur Verfügung. Hiervon entfallen 61 Datensätze auf die HA- und 63 Datensätze auf die LA-Bedingung.

Vor der Analyse der Kraft, die beim Klicken auf die Maus ausgeübt wurde, müssen zunächst die Klickzeitpunkte detektiert werden. Die Vorgehensweise bei der Extraktion der Mausklicks aus dem kontinuierlichen Kraftsignal ist in Abschnitt 5.2.1.3 beschrieben.

Wie aus Abbildung 5.10 ersichtlich wird, gibt es bei der auf die Maustaste ausgeübten Kraft große Unterschiede. Dies ist zum einen auf die interindividuellen Unterschiede beim Klicken der Maustaste zurückzuführen als auch auf Kalibrierungsunterschiede zwischen den einzelnen Mäusen. Das abgebildete Streudiagramm zeigt, wie groß die durchschnittliche auf die

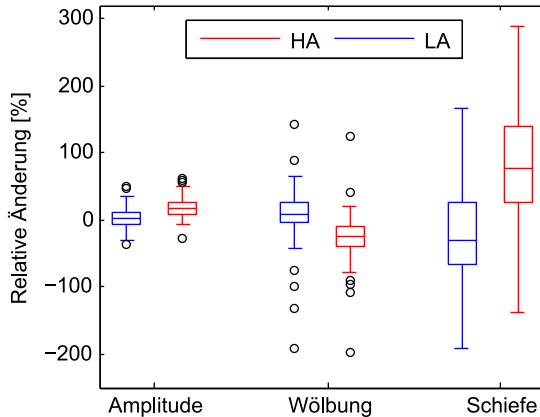




**Abbildung 5.10:** Streuung der Klickkraft während des Spiels im Vergleich zur Baseline für beide experimentellen Bedingungen

Maustaste ausgeübte Kraft während des Spiels im Vergleich zur Baseline war. Anhand der Regressionsgeraden lässt sich gut erkennen, dass in der HA-Bedingung während des Spiels eine größere Krafteinwirkung stattfindet als während der LA-Bedingung.

Aufgrund der großen interindividuellen Unterschiede bei der auf die Maustaste ausgeübten Kraft, wird im Folgenden die relative Kraftänderung betrachtet. Da während der Baseline-Phase keine Mausaktivität erfolgt, wird zur Normierung die Klickkraft außerhalb des Spiels verwendet. In Abbildung 5.11 sind die Unterschiede der Mausklicks während des Spiels in Bezug auf die ausgeübte Kraft, die Wölbung sowie die Schiefe beider experimentellen Bedingungen dargestellt. Es zeigt sich, dass unter hoher emotionaler Erregung eine größere Kraft auf die Maustaste ausgeübt wird (+19,55%), als bei niedriger emotionaler Erregung, bei welcher die Kraft mit nur 3,38% nahezu konstant bleibt. Zudem findet eine deutliche Veränderung der Morphologie der Mausklicks statt, was sich an Wölbung und Schiefe zeigt: Die durchschnittliche Wölbung nimmt unter hoher Erregung



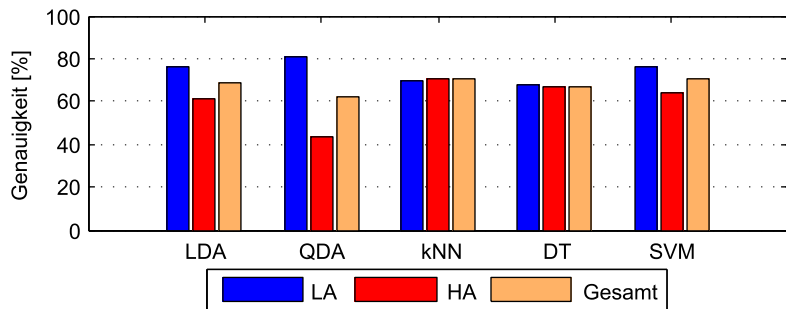
**Abbildung 5.11:** Relative Veränderung der Mausclicks während des Spiels

deutlich ab ( $-34,64\%$ ), während sich die Schiefe unter hoher Erregung mit einem durchschnittlichen Anstieg um  $109,87\%$  mehr als verdoppelt. Dies kann damit erklärt werden, dass das Klicken unter hoher Erregung üblicherweise schneller erfolgt als bei niedriger emotionaler Erregung.

Ein Vergleich der verschiedenen Klassifikatoren zeigt, dass die Klassifikationsgenauigkeiten für LDA, kNN und SVM recht dicht beieinander liegen (LDA:  $68,83\%$ , kNN:  $70,11\%$ , SVM:  $70,21\%$ ). Die SFS liefert den dreidimensionalen Featurevektor bestehend aus Amplitude, Wölbung und Schiefe der Mausclicks als optimale Featurekombination was der Summe der berechneten Features entspricht. In Abbildung 5.12 ist der Vergleich aller Klassifikatoren dargestellt.

## 5.2.6 Multimodale Klassifikation

Nach dem Aufbau der Klassifikatoren für die einzelnen Signale erfolgt im letzten Schritt die Kombination der Ergebnisse der einzelnen Klassifikatoren. Zu diesem Zweck kommt majority voting zum Einsatz, da dieses Verfahren unabhängig von der Anzahl der verfügbaren Klassifikationsergebnisse eingesetzt werden kann. Insbesondere unter Alltagsbedingungen



**Abbildung 5.12:** Klassifikationsgenauigkeiten für emotionale Erregung basierend auf den mausbasierten Features

bringt dieses Verfahren den Vorteil mit sich, dass auch bei Ausfall eines einzelnen Signals weiterhin eine Klassifikation basierend auf den verbleibenden Signalen erfolgen kann. Zwei verschiedene Verfahren werden des majority votings werden verglichen: gewichtetes und ungewichtetes majority voting. Beim gewichteten Voting gehen die Ergebnisse der Klassifikatoren gewichtet mit ihrer ex post-Wahrscheinlichkeit in das Voting ein. Beim ungewichteten Voting hingegen ist die Anzahl der Votes für eine Klasse entscheidend. Lediglich wenn gleich viele Votes für zwei Klassen vorliegen, wird die ex post-Wahrscheinlichkeit mit in die Entscheidung einbezogen. Bei der Kombination von nur zwei Klassifikatoren liefern beide Verfahren somit das gleiche Ergebnis, während es bei der Kombination mehrerer Klassifikatoren zu Unterschieden kommen kann.

In Tabelle 5.10 sind noch einmal alle Klassifikationsgenauigkeiten der zuvor aufgebauten Klassifikatoren auf den einzelnen physiologischen Signalen dargestellt. Es zeigt sich, dass für alle Klassifikatoren mittels der kardiovaskulären Signale die besten Klassifikationsgenauigkeiten erzielt werden können während die Klassifikationsgenauigkeiten für die EDA-basierte Klassifikation bei nahezu allen Klassifikatoren dicht am Zufallsniveau liegen.

Tabelle 5.11 zeigt die Ergebnisse der Klassifikation für unterschiedliche Signalkombinationen. Für alle verwendeten Klassifikatoren mit Ausnahme der QDA zeigt sich durch Kombination verschiedener physiologischer

	Klassifikationsgenauigkeit [%]				
	SVM	QDA	LDA	kNN	DT
EKG	<b>80,91</b>	79,66	80,69	78,47	73,10
PPG	80,12	<b>80,25</b>	<b>81,53</b>	<b>78,75</b>	<b>74,18</b>
Maus	70,21	62,35	68,83	70,11	66,61
EDA <sub>zeit</sub>	59,03	48,13	56,53	45,76	47,50
EDA <sub>event</sub>	47,86	52,82	53,42	47,33	45,51

**Tabelle 5.10:** Klassifikationsgenauigkeiten der verschiedenen Klassifikatoren für einzelne physiologische Signale

Signale eine Verbesserung des Klassifikationsergebnisses. Jedoch liegt die Verbesserung lediglich im Bereich von etwa 2%. Es kann somit festgehalten werden, dass die Kombination mehrerer physiologischer Signale zwar in der Regel zu einer Verbesserung der Klassifikationsgenauigkeit führt, diese jedoch verhältnismäßig gering ausfällt. Dies deckt sich auch mit den Ergebnissen von D’Mello & Kory (2012). Zusätzlich zeigen die Ergebnisse, dass die Hinzunahme von EDA zur multimodalen Klassifikation in allen Fällen zu einer Verschlechterung des Gesamt-Klassifikationsergebnisses führt.

	Klassifikationsgenauigkeit [%]				
	SVM	QDA	LDA	kNN	DT
EKG & PPG	<b>82,69</b>	<b>79,66</b>	82,80	80,18	<b>75,95</b>
EKG & Maus	80,41	77,66	80,41	79,96	75,24
PPG & Maus	77,61	62,35	77,27	75,78	72,43
EDA <sub>time</sub> & EDA <sub>event</sub>	48,58	38,76	51,68	48,79	52,16
EKG&PPG&Maus (count)	81,72	77,66	81,88	80,10	75,85
EKG&PPG&Maus (weight)	-	77,66	<b>83,23</b>	<b>80,31</b>	75,13
Alle Signale (count)	80,38	76,52	81,48	77,67	73,29
Alle Signale (weight)	-	75,86	81,55	78,09	72,95

**Tabelle 5.11:** Klassifikationsgenauigkeiten der verschiedenen Klassifikatoren für unterschiedliche Signalkombinationen

Ein Laufzeitvergleich der einzelnen Klassifikatoren zeigt, dass es hier deutliche Unterschiede gibt: Während die Laufzeit von LDA und kNN nahezu gleich ist, dauert die Klassifikation mittels SVMs etwa 148 mal länger als mit den anderen Verfahren, weshalb den einfacheren Verfahren

der Vorzug gegeben werden sollte. Basierend auf den vorliegenden Daten kristallisiert sich somit die LDA sowohl hinsichtlich Laufzeit als auch in Bezug auf die Klassifikationsgenauigkeit als das zu bevorzugende Verfahren heraus.

### 5.3 Übertragbarkeit der Ergebnisse

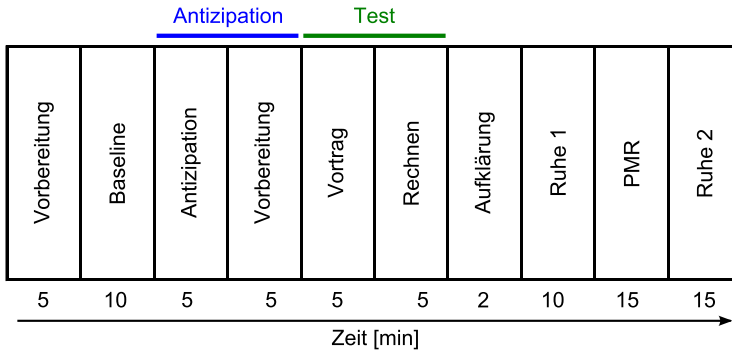
Zur Überprüfung, ob die Ergebnisse der oben dargestellten Analysen auf andere Studien übertragbar sind, wurden sie anhand einer weiteren Studie zur emotionalen Erregung überprüft. Des Weiteren wurde anhand dieser Studie untersucht, wie gut die Klassifikation auf einem Mehrklassenproblem umgesetzt werden kann.

#### 5.3.1 Studiendesign

Für die Erfassung der physiologischen Daten im Rahmen dieser Studie wurde das SOMNOscreen Messsystem von Somnomedics verwendet. Dieser mobile Polygraph ermöglicht die synchrone Messung der physiologischen Parameter EKG, EDA, PPG sowie der Atmung. Zusätzlich zu den physiologischen Daten erfolgte nach jeder Phase der Studie eine Messung des Cortisol-Spiegels der Studienteilnehmer. Da im Rahmen dieser Studie keine Eingaben über die Maus erfolgten, konnte die kraftsensitive Maus nicht eingesetzt werden.

An der Studie nahmen insgesamt 30 Versuchspersonen teil, jeweils zu gleichen Teilen männlich und weiblich. Das Durchschnittsalter der Versuchspersonen lag bei 24 Jahren mit einer Standardabweichung von 3 Jahren. Zur Induktion der emotionalen Erregung wurde der Trier Social Stress Test (TSST) verwendet (Kirschbaum et al., 1993). Der TSST besteht im Wesentlichen aus zwei je 10-minütigen Phasen: einer Antizipations- und einer Testphase. Während der Antizipationsphase wird zunächst die Vorgehensweise erklärt. Der Teilnehmer soll sich in die Rolle eines Jobbewerbers versetzen und einen 5-minütigen Vortrag vor einem dreiköpfigen Auswahlkomitee halten. Im Anschluss daran erhält er 5 Minuten Zeit, sich vorzubereiten. In der Testphase muss der Teilnehmer zunächst seinen Vortrag halten. Ist er früher fertig, werden von dem Auswahlkomitee weitere Fragen gestellt. Nach Abschluss dieser Phase erhält der Teilnehmer

Informationen über die nächste Aufgabe, welche darin besteht, von einer vierstelligen Zahl in 13er Schritten rückwärts zu zählen. Bei jedem fehlerhaften Ergebnis muss von vorne begonnen werden. Während des ganzen Tests läuft scheinbar eine Videokamera mit. Im Anschluss an den Test erfolgt ein Debriefing, bei welchem der Teilnehmer über den Inhalt der Studie aufgeklärt wird.



**Abbildung 5.13:** Zeitlicher Ablauf der Studie und Einteilung der Phasen in unterschiedliche Erregungsniveaus

Im Rahmen dieser Studie wurde der Test zusätzlich um eine Baseline-Phase vor dem Test sowie weitere Ruhephasen im Anschluss an den Test erweitert. Der genaue Studienablauf ist in Abbildung 5.13 dargestellt.

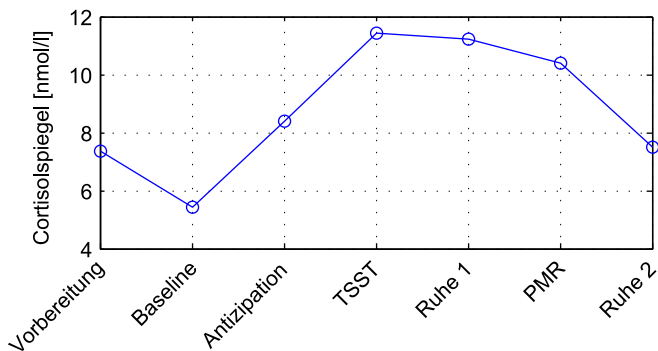
Im Anschluss an die jeweiligen Phasen erfolgte jeweils eine Bestimmung des Cortisol-Spiegels sowie eine Befragung zur Befindlichkeit mittels des Multidimensionalen Befindlichkeitsfragebogens (MDBF).

### 5.3.2 Datenanalyse

Im ersten Schritt der Datenanalyse wird anhand der Cortisolmessungen sowie der Ergebnisse des MDBF untersucht, welche Erregungsniveaus in den einzelnen Phasen des Experiments vorlagen. Aus diesen werden unterschiedliche Stufen für die anschließende Klassifikation abgeleitet. Im nächsten Schritt erfolgt eine Klassifikation der Daten nach dem in den vorhergehenden Abschnitten vorgestellten Klassifikationskonzept.

### 5.3.2.1 Analyse der Erregungsniveaus

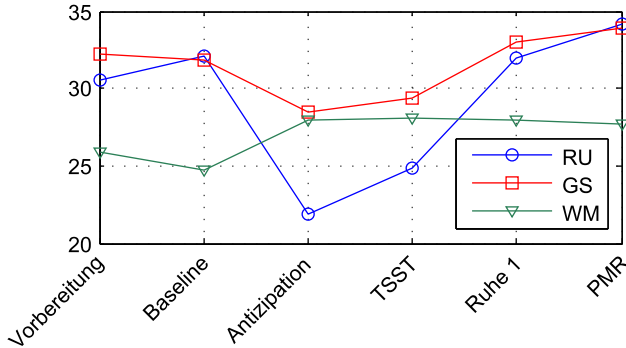
Wie bereits bei der vorstehend beschriebenen Studie, soll auch im Rahmen dieser Studie zunächst überprüft werden, inwiefern durch das Experiment Erregung induziert wurde. Hierfür werden die Cortisol-Messungen nach den jeweiligen Phasen des Experiments sowie der MDBF herangezogen. In Abbildung 5.14 ist der durchschnittliche Verlauf des Cortisolspiegels für die einzelnen Phasen des Experiments dargestellt. Da die Cortisol-Reaktion üblicherweise 10 bis 20 Minuten verzögert erfolgt, wurde eine entsprechende Korrektur vorgenommen, um die Messungen den einzelnen Phasen des Experiments zuordnen zu können. Der Verlauf des Cortisolspiegels zeigt, dass für die Phasen des TSST ein deutlicher Anstieg zu beobachten ist. Der Cortisolabbau im Körper erfolgt nur sehr langsam, daher treten auch nach dem TSST noch relativ hohe Werte auf. Ein stetiges Absinken des Cortisolspiegels nach dem TSST lässt jedoch auf ein zunehmendes Sinken des Erregungsniveaus schließen.



**Abbildung 5.14:** Mittlerer Cortisolspiegel über alle Versuchspersonen für die einzelnen Phasen des Experiments

Auch der MDBF bestätigt die Beobachtungen bzgl. des Erregungsniveaus. Abbildung 5.15 zeigt den Verlauf für die drei Dimensionen Ruhe – Unruhe (RU), Gut – Schlecht (GS) sowie Wach – Müde (WM). Je höher ein Wert dabei ausfällt, desto positiver dessen Ausprägung. Für die Beurteilung des Erregungsniveaus ist vor allem die Dimension RU von Bedeutung, welche in blau dargestellt ist. Hohe Werte lassen hierbei auf ein niedriges

Erregungsniveau schließen, während bei hoher Erregung niedrigere Werte auftreten. Auch in dieser Befragung zeigt sich deutlich, dass der TSST zu einer deutlichen Steigerung des Erregungsniveaus führt. Im Gegensatz zum Cortisolspiegel zeigt sich anhand des MDBF auch deutlich das Abfallen des Erregungsniveaus im Anschluss an den Test.



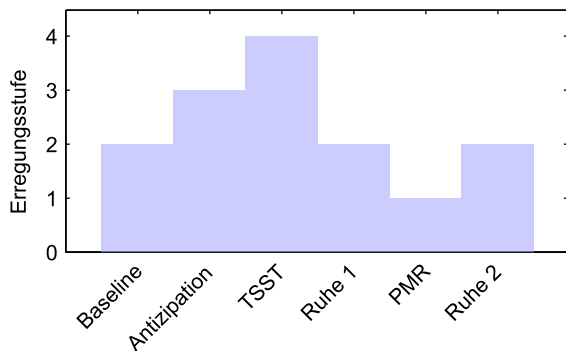
**Abbildung 5.15:** Mittlere MDBF-Werte über alle Versuchspersonen für die einzelnen Phasen des Experiments (RU: Ruhe – Unruhe, GS: Gut – Schlecht, WM: Wach – Müde)

Basierend auf den Ergebnissen der Analyse des Cortisolspiegels sowie des MDBFs wird eine Aufteilung der einzelnen Phasen des Experiments in vier unterschiedliche Erregungsniveaus vorgenommen. Diese sind in Abbildung 5.16 dargestellt. Die Phasen der Vorbereitung sowie der Aufklärung über den Hintergrund des Experiments werden bei den weiteren Analysen nicht weiter in Betracht gezogen, da aufgrund der Datenlage keine eindeutigen Tendenzen bezüglich des Erregungsniveaus erkennbar waren.

### 5.3.2.2 Klassifikation der emotionalen Erregung

Wie in Abbildung 5.16 dargestellt, kann die Studie in insgesamt vier verschiedene Erregungsniveaus unterteilt werden. Zur besseren Vergleichbarkeit erfolgt die Klassifikation zunächst auf einem Zwei-Klassenproblem und wird im nächsten Schritt auf ein Vier-Klassenproblem erweitert. Erneut werden alle Analysen basierend auf einer Fensterlänge von 15 Sekunden durchgeführt.





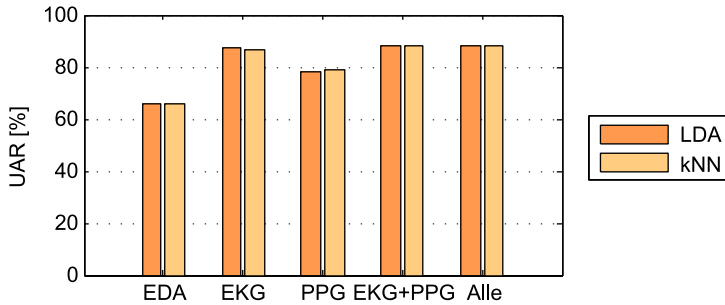
**Abbildung 5.16:** Einteilung der experimentellen Phasen in unterschiedliche Erregungsstufen

Da sich aufgrund der unterschiedlichen Längen der einzelnen Phasen die Anzahl der Samples in den einzelnen Klassen deutlich unterscheidet, wird zum Vergleich der Klassifikationsgenauigkeit der UAR herangezogen.

Bei der Klassifikation zweier Klassen werden die Daten der Baseline, der Ruhephasen sowie der PMR zu der Klasse der niedrigen Erregung zusammengefasst. Die Daten des TSST, bestehend aus Antizipations- und Testphase, bilden die Klasse der hohen Erregung. Zur Klassifikation werden LDA und kNN ( $k = 13$ ) verglichen, da diese zusammen mit den SVMs in der vorherigen Studie die besten Ergebnisse geliefert haben. SVMs werden aufgrund der schlechten Laufzeiteigenschaften nicht weiter betrachtet.

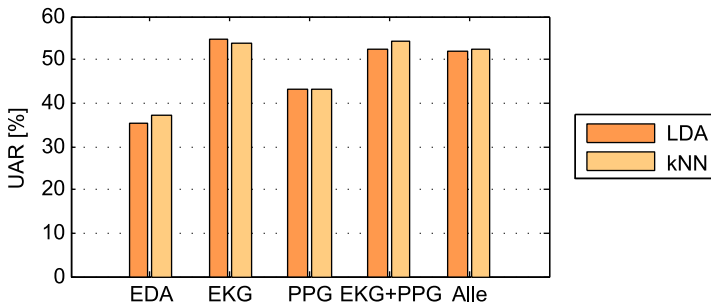
In Abbildung 5.17 sind die Ergebnisse der Klassifikation für zwei Klassen dargestellt. Es zeigt sich, dass auch hier EDA mit Abstand am schlechtesten zur Trennung unterschiedlicher Erregungszustände geeignet ist. Die besten Klassifikationsergebnisse hingegen können auch hier mit einer Kombination der EKG- und PPG-Klassifikatoren erzielt werden. Eine Kombination aller Signale führt wiederum zu einer leichten Verschlechterung der Klassifikationsgenauigkeit. Hinsichtlich des Gesamtergebnisses schneidet kNN mit einem UAR von 88,30 % minimal besser ab als die LDA.

Im nächsten Schritt werden die Daten als vier-Klassen-Problem analysiert. Die Klassenbildung hierfür erfolgt gemäß der in Abbildung 5.16 dargestellten Unterteilung. Die Ergebnisse sind in Abbildung 5.18 darge-



**Abbildung 5.17:** Gesamtklassifikationsergebnisse bei zwei Klassen für die einzelnen physiologischen Signale sowie deren Kombination

stellt. Insgesamt ist im Vergleich zum zwei-Klassen-Problem ein deutlicher Rückgang der Klassifikationsgenauigkeit zu erkennen. Jedoch ist dieses immer noch deutlich vom Zufallsniveau entfernt. Es zeigt sich erneut, dass EDA zur Unterscheidung der Erregungszustände am wenigsten geeignet ist: der UAR liegt bei lediglich 35,08 % für kNN bzw. 35,31 % für die LDA. Das beste Ergebnis findet sich auch hier bei der Kombination des EKG- und des PPG-Klassifikators. Für kNN kann hier ein UAR von 54,10 %, für die LDA ein UAR von 52,42 % erzielt werden.



**Abbildung 5.18:** Klassifikationsergebnisse bei vier Klassen für die einzelnen physiologischen Signale sowie deren Kombination

Das Ranking der einzelnen Features hinsichtlich ihrer Diskriminierungsfähigkeit in der hier untersuchten Studie fällt ähnlich aus, wie bereits

bei der vorhergehenden Studie. Bei der EDA lassen das SCL sowie die Hjorth-Parameter am ehesten Rückschlüssen über das Erregungsniveau einer Person zu. Für den EKG-Klassifikator stellen die mittlere Herzrate, RMSSD sowie SD1 erneut die besten 3 Features zur Klassifikation dar. Ähnlich verhält es sich auch für den PPG-Klassifikator. Das zuverlässigste System kann basierend auf der Kombination aus EKG- und PPG-Klassifikator erstellt werden.

Zusammenfassend kann festgestellt werden, dass das aufgebaute Gesamtsystem auch auf Daten aus anderen Studien übertragen werden kann. Ferner kann auch eine Erweiterung auf mehr als zwei Klassen als unproblematisch angesehen werden.

### 5.4 Zusammenfassung der Ergebnisse

In diesem Kapitel wurde eine Studie vorgestellt, bei welcher untersucht wurde, welche physiologischen Signale am besten geeignet sind, um emotionale Erregung abzubilden. Hierzu wurde eine Studie mit einem Zwischensubjekt-Design vorgestellt, bei welcher mittels eines Spiels niedrige und hohe emotionale Erregung induziert wurden. Untersucht wurden hierbei EKG, PPG, EDA sowie die auf die Maustaste ausgeübte Kraft. Zudem erfolgte eine Validierung der Ergebnisse anhand der Daten aus einer anderen Studie zur emotionalen Erregung. Anhand der Daten der ersten Studie zeigte sich, dass EDA eher zur Klassifikation kurzfristiger Erregungsänderungen als zur Klassifikation längerandauernder Erregungszustände geeignet ist. Für EKG, PPG sowie die Maustastenkraft konnten jedoch in der unimodalen Klassifikation Erkennungsraten zwischen 68,83 % und 81,07 % erzielt werden. Durch eine Kombination der Klassifikatoren konnte die Klassifikationsgenauigkeit auf 83,32 % erhöht werden. Die besten Klassifikationsergebnisse konnten mittels LDA erzielt werden. Der Performancegewinn durch die Verwendung eines multimodalen Klassifikators im Vergleich zur unimodalen Klassifikation liegt somit insgesamt bei 2,77 %. Dieses Ergebnis ähnelt den von D'Mello & Kory (2012) analysierten Studien. Bei dieser Metaanalyse konnte gezeigt werden, dass bei Studien mit natürlichen bzw. seminaturalen Emotionen (im Gegensatz zu gespielten Emotionen) mittels multimodaler Klassifikation im Durchschnitt lediglich ein Performancezuwachs von 4,86 % im Vergleich zur unimodalen

Klassifikation erzielt werden konnte.

In der Validierungsstudie zeigten sich ähnliche Ergebnisse. Auch hier erzielte die unimodale Klassifikation für EDA mit Abstand die schlechtesten Ergebnisse. Durch Kombination der unterschiedlichen Klassifikatoren konnte für zwei Klassen ein maximaler UAR von 88,83 % mittels LDA erzielt werden. Bei der Analyse von vier Klassen lag der maximale UAR bei 54,10 %. In der Validierungsstudie konnten mittels kNN minimal bessere Ergebnisse gegenüber der LDA erzielt werden.

Die Ergebnisse der Klassifikation können als sehr gut bewertet werden, da in den vorstehenden Analysen ein personeneunabhängiger Klassifikator verwendet wurde. Zwar gibt es Studien, bei denen eine Klassifikationsgenauigkeit von über 90 % berichtet wurde, jedoch konnten derart hohe Genauigkeiten nur mit personenabhängiger Klassifikation erreicht werden.

Ein großer Vorteil des vorgestellten Klassifikators ist darin zu sehen, dass dieser aufgrund des kaskadierenden Aufbaus auch dann noch in der Lage ist, valide Ergebnisse zu liefern, wenn einzelne Signale kurzzeitig ausfallen. Dies ist insbesondere unter Alltagsbedingungen ein wichtiger Aspekt, da es oftmals vorkommt, dass einzelne Signale aufgrund von Artefakten wie abgefallenen Elektroden, leerer Batterien des Systems oder Ähnlichem gestört sind.



# 6 Eignung von EDA für die kurzfristige Erregungserkennung

In den in Kapitel 5 durchgeführten Analysen hat sich gezeigt, dass EDA zur Klassifikation langfristiger Erregungszustände nur bedingt geeignet ist. Jedoch zeichnet sich an den in Abschnitt 5.2.4 dargestellten Ergebnissen eine gute Eignung von EDA für die Erkennung kurzfristiger Erregungsänderungen ab. Diese Fragestellung soll daher in dem folgenden Kapitel genauer untersucht werden.

Kurzfristige Änderungen des emotionalen Erregungsniveaus werden in der Regel durch das Auftreten entsprechender emotionaler Stimuli ausgelöst. Im Gegensatz zu Laborumgebungen sind die Auftretenszeitpunkte emotionaler Stimuli unter Alltagsumgebungen in der Regel unbekannt. Hinzu kommt, dass es bisher nur wenige Untersuchungen bezüglich der Robustheit des EDA-Signals unter Alltagsbedingungen gibt. Aus diesem Grund wird zusätzlich der Einfluss von Bewegung auf das EDA-Signal in die Untersuchungen mit einbezogen.

## 6.1 Studiendesign

Ziel der nachfolgend dargestellten Studie ist es, herauszufinden, wie emotionale Stimuli in einem EDA-Signal möglichst akkurat identifiziert werden können und inwiefern Kontextinformationen Aufschluss darüber geben können, in welchen Situationen keine Nutzung des EDA-Signals mehr möglich ist. Studienaufbau und Ablauf sind nachfolgend dargestellt.

### 6.1.1 Studienaufbau

Nachfolgend sind die Hard- und Softwarelösungen beschrieben, welche im Rahmen der Studie zum Einsatz kamen.

### 6.1.1.1 Hardware

Einen wichtigen Aspekt der Studie stellte die Analyse von EDA-Signalen hinsichtlich der Erkennung emotionaler Einflüsse in mobilen Szenarien dar. Aus diesem Grund wurde der edaMove-Sensor für die Aufzeichnung der EDA-Daten verwendet, da dieser im Vergleich zu anderen Systemen relativ belastungsfrei getragen werden kann und zugleich eine hohe Signalqualität bei der Aufzeichnung ermöglicht. Die Elektroden werden bei diesem System an der Hand platziert. Des Weiteren kam der ekgMove-Sensor zum Einsatz, da das EKG für weitere Kontextinformationen herangezogen werden kann. Beide Systeme verfügen zusätzlich über einen integrierten 3D-Beschleunigungssensor. Eine Beschreibung beider Systeme findet sich in Abschnitt 4.2.1.

### 6.1.1.2 Software

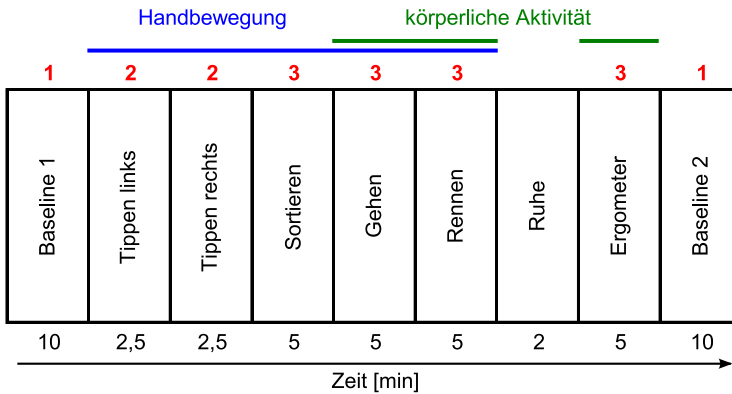
Zur Aufzeichnung der Daten wurde der movisens SensorManager verwendet, welcher ein synchrones Starten der Messungen auf den verwendeten Sensoren ermöglicht. Da im Rahmen dieser Studie keine Online-Analyse der Daten erforderlich war, wurde auf den Einsatz der xAffect-Software verzichtet.

Zur Steuerung des Studienablaufs wurde Microsoft PowerPoint genutzt. Mittels einer automatisierten PowerPoint-Präsentation wurden die Versuchspersonen durch den kompletten Studienablauf geführt.

## 6.1.2 Studienablauf

An der Studie nahmen insgesamt 31 Personen teil. 22 der Versuchspersonen waren männlich, 9 weiblich. Das Durchschnittsalter lag bei 23,5 Jahren. Alle Messungen wurden in den Räumen des FZI durchgeführt.

Der genaue Studienablauf ist in Abbildung 6.1 schematisch dargestellt. Zu Beginn und Ende der Studie fand jeweils eine zehnmütige Baseline statt. Im Anschluss an die Baseline folgten Phasen mit verschiedenen Handaktivitäten und Bewegungsniveaus. Die Phasen des Tippens – welches abwechselnd zunächst mit der linken und anschließend mit der rechten Hand erfolgte – und des Karten Sortierens dienten der Untersuchung des Einflusses von Handbewegung auf das an der Hand gemessene EDA-Signal.



**Abbildung 6.1:** Zeitlicher Ablauf der Studie. Rote Zahlen geben die Anzahl der emotionalen Stimuli in jeder Phase an

Zusätzlich wurden unterschiedliche Niveaus der körperlichen Anstrengung untersucht: Beim Gehen und Rennen wurde mittlere bzw. starke Bewegung mit mittlerer bzw. starker körperlicher Anstrengung kombiniert. Während des Fahrens auf einem Ergometer wurde körperliche Anstrengung ohne Bewegung hervorgerufen. In dieser Phase wurden die Versuchspersonen zusätzlich angewiesen, ihre Hände nicht zu bewegen. Zur Minimierung des Einflusses der einzelnen Phasen auf die jeweils nächste Phase erfolgte die Anordnung der Phasen beginnend mit ruhigeren Aktivitäten hin zu anstrengenderen Aktivitäten. Zusätzlich gibt es eine kurze Ruhephase zwischen Rennen und Ergometer fahren.

Während der jeweiligen Phasen wurden den Versuchspersonen mehrmals gleichzeitig ein visueller und ein akustischer Reiz präsentiert, welcher der Induktion emotionaler Reaktionen diente. Hierzu wurde das Bild einer Horrorgestalt in Kombination mit einem Frauenschrei verwendet. Zur besseren Synchronisierbarkeit der Daten der einzelnen Versuchspersonen, erfolgte die Präsentation der emotionalen Stimuli für jede Versuchsperson nach dem gleichen Zeitschema.



## 6.2 Datenanalyse

Nachfolgend ist die Analyse der aufgezeichneten Daten dargestellt. Während in Abschnitt 6.2.2 auf die Erkennung der elektrodermalen Reaktionen für die einzelnen Versuchspersonen eingegangen wird, erfolgt in Abschnitt 6.2.3 eine übergreifende Analyse der Daten über alle Versuchspersonen.

Aufgrund von Störungen im Versuchsablauf mussten zwei Datensätze aus den Analysen ausgeschlossen werden. Bei einer Versuchsperson kam es während der Aufnahmen zu Problemen beim Elektrode-Haut-Kontakt während bei der anderen Versuchsperson Probleme bei der Induktion der akustischen Stimuli auftraten.

### 6.2.1 Vorbereitung der Daten

Da es bei den EDA-Messungen teilweise zu sehr großen interindividuellen Unterschieden hinsichtlich des Wertebereichs des Signals kam, wurde bei einigen Analysen eine Normierung des Signals durchgeführt. In Anlehnung an Schestatsky et al. (2010) wurde hierfür eine Division des Gesamtsignals durch den Mittelwert der ersten Baseline durchgeführt. Zudem wurde eine Trennung des Signals in den phasischen und den tonischen Anteil vorgenommen. Zur Extraktion des phasischen Anteils wurde ein Butterworth-Hochpassfilter 2. Ordnung mit einer Grenzfrequenz von 0,2 Hz verwendet, bei der tonischen Aktivität ein Tiefpassfilter mit einer Grenzfrequenz von 0,08 Hz.

Aus dem in den edaMove-Sensor integrierten 3D-Beschleunigungssensor wird die Aktivität wie folgt berechnet:

$$a(t) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2} \quad (6.1)$$

Zur Kompensation von Einflüssen der Erdbeschleunigung wird vor der Aktivitätsberechnung eine Hochpassfilterung der einzelnen Beschleunigungssignale mit einem Butterworth-Filter 2. Ordnung und einer Grenzfrequenz von 0,25 Hz durchgeführt.

## 6.2.2 Erkennung von EDA-Reaktionen

Im ersten Schritt der Analyse soll die Fragestellung untersucht werden, mit welcher Genauigkeit die EDA-Reaktionen in den einzelnen Phasen des Experiments identifizierbar sind.

Drei verschiedene Verfahren der Detektion der EDA-Reaktionen werden hierbei verglichen:

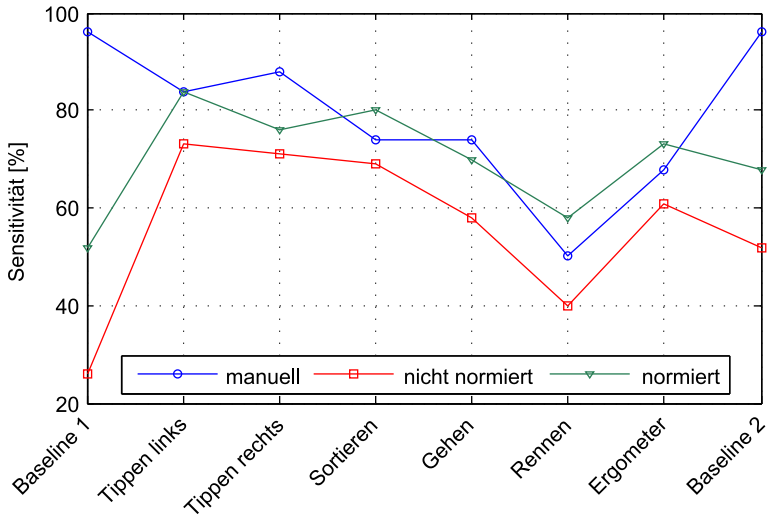
- Manuelle Detektion
- Computergestützte Detektion ohne Normierung
- Computergestützte Detektion mit Normierung

Im Rahmen der manuellen Detektion werden die einzelnen Signale gesichtet und die Reaktionen ohne Kenntnis über die genauen Stimuluszeitpunkte annotiert.

Die EDA-Reaktionen wurden mittels des in Abschnitt 2.3.1.2 dargestellten Algorithmus detektiert. Da bei dem nicht-normalisierten Signal die Reaktionen oft weitaus höher als die in Dawson et al. (2007) spezifizierten Werte ausfielen und teilweise Amplituden von über  $5\ \mu\text{S}$  auftraten, wurde für die Plausibilitätsprüfung der Wertebereich für die zulässige Amplitudenhöhe auf  $0,5\ \mu\text{S}$  bis  $3\ \mu\text{S}$  erweitert, da mit diesem Bereich bei der Erkennung die besten Ergebnisse erzielt werden konnten. Bei dem normalisierten Signal wurden die Grenzen bei  $0,05\ \mu\text{S}$  und  $1,5\ \mu\text{S}$  belassen.

In Abbildung 6.2 ist die Sensitivität für die Detektion der einzelnen EDA-Reaktionen dargestellt. Diese spiegelt wider, wie viele der emotionalen Events korrekt detektiert werden konnten. Es ist deutlich zu erkennen, dass die Normierung des Signals zu einer deutlichen Steigerung der Sensitivität führt. Jedoch ist die manuelle Detektion der Reaktionen der computergestützten Detektion oftmals überlegen. Dies kann vor allem darauf zurückgeführt werden, dass die emotionalen Reaktionen in der Baseline bei vielen Versuchspersonen sehr hoch ausfallen und somit außerhalb des Wertebereiches bei der Plausibilitätskontrolle liegen. Jedoch führt eine Veränderung des Wertebereiches dahingehend, dass auch Reaktionen mit größerer Amplitudenhöhe berücksichtigt werden, nicht zu einer Verbesserung der Erkennung, da dies auch die Anzahl der falsch detektierten Reaktionen erhöht.

Die Ergebnisse zeigen, dass die Sensitivität für Phasen, in denen lediglich

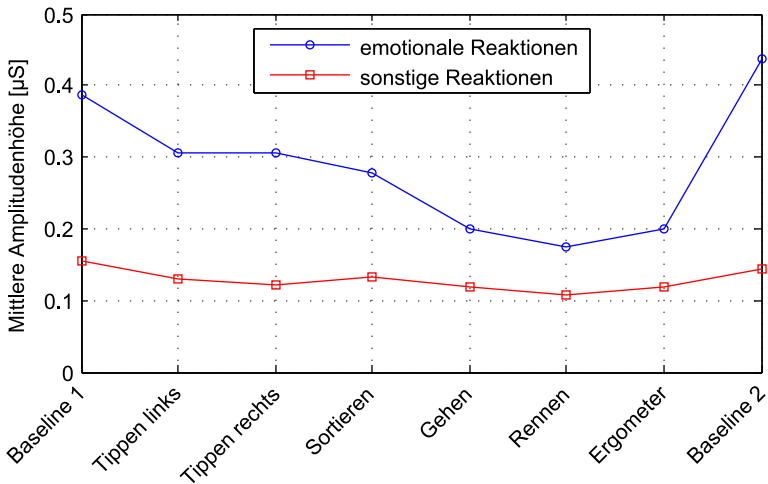


**Abbildung 6.2:** Sensitivität verschiedener Methoden zur Detektion von EDA-Reaktionen

die Hände genutzt werden, weitaus höher ist, als für Phasen hoher körperlicher Aktivität. Überraschend sind die relativ geringen Unterschiede für das Tippen mit der rechten bzw. linken Hand. Dies kann als Indikator dafür gesehen werden, dass eine Messung der EDA an der Hand in Szenarien, bei welchen ein Computer ohne körperliche Anstrengung bedient wird, durchaus geeignet ist. Je höher jedoch die körperliche Anstrengung, desto mehr Fehldetektionen treten auf. Selbst für die manuelle Detektion sinkt die Sensitivität auf 50 %.

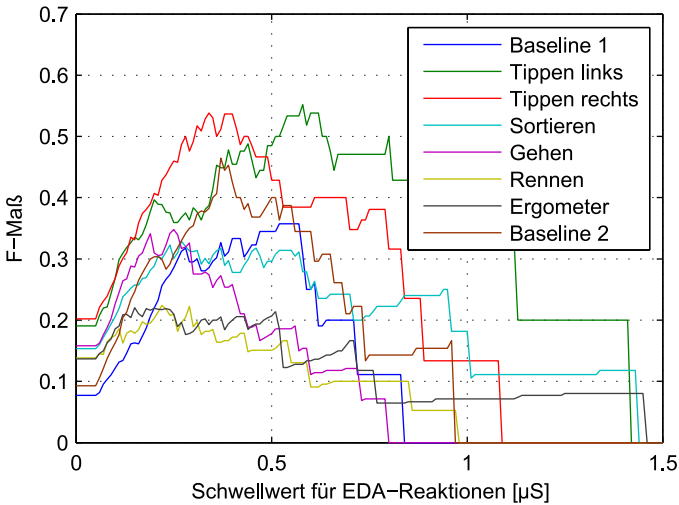
Abbildung 6.3 zeigt die mittlere Amplitudenhöhe der detektierten EDA-Reaktionen. Hierbei wird getrennt zwischen den Reaktionen, welche in Verbindung mit einem emotionalen Stimulus auftreten, sowie allen weiteren detektierten Reaktionen. Es wird ersichtlich, dass die Reaktionen, welche in Verbindung mit einem emotionalen Stimulus stehen, deutlich höher ausfallen, als diejenigen, welche auf anderen Ursachen wie beispielsweise Bewegung basieren. Zusätzlich zeigt sich, dass die Amplitudenhöhe der emotionalen Reaktionen bei steigender körperlicher Anstrengung deutlich

zurückgeht. Eine mögliche Erklärung hierfür kann darin gesehen werden, dass die emotionalen Reaktionen bei körperlicher Aktivität zunehmend von thermoregulatorischen Aktivitäten der Schweißdrüsen überlagert werden und hierdurch Sättigungseffekte eintreten. Für die Trennung zwischen emotional ausgelöst sowie sonstigen EDA-Reaktionen bedeutet dies, dass die Trennung dieser beiden Gruppen mit zunehmender körperlicher Aktivität schwieriger wird.



**Abbildung 6.3:** Vergleich der mittleren Amplitudenhöhe emotionaler und nicht-emotionaler EDA-Reaktionen

Zur Bestimmung eines geeigneten Schwellwertes für die Trennung der emotional ausgelöst EDA-Reaktionen von den anderen Reaktionen soll untersucht werden, wie der Schwellwert für die Zuordnung der Reaktionen optimal gesetzt sein sollte, damit ein möglichst großer Anteil der emotionalen Stimuli als solche detektiert wird. Zur Ermittlung der optimalen Trennung der Amplituden wird das F-Maß herangezogen, da dieses eine Kombination aus Sensitivität und positiver Prädiktivität darstellt. Abbildung 6.4 zeigt das F-Maß für die einzelnen Phasen des Experiments in Abhängigkeit des Schwellwertes. Wie bereits angenommen, fällt die Trennung emotionaler EDA-Reaktionen von sonstigen Reaktionen für die Phasen mit hoher körperlicher Anstrengung am schwersten. Der maximal



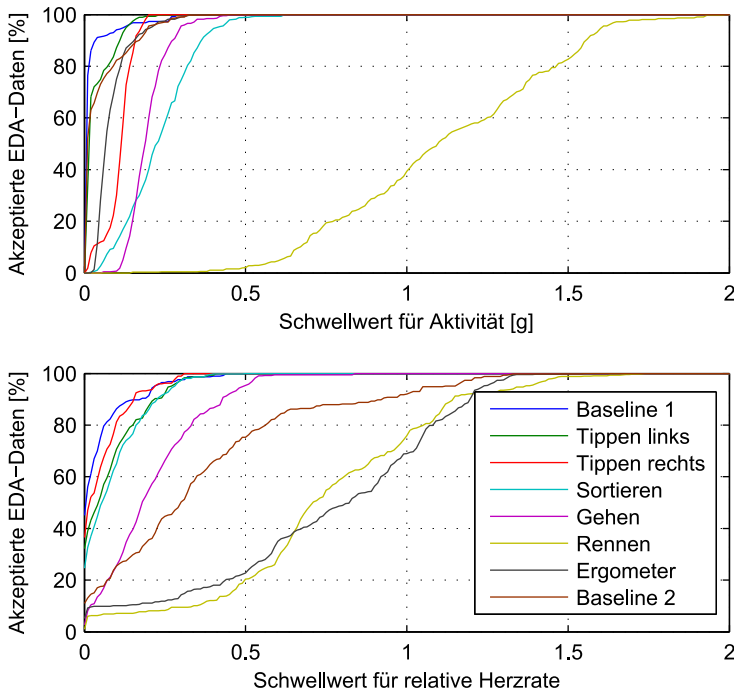
**Abbildung 6.4:** F-Wert für die Detektion emotionaler EDA-Reaktionen in Abhängigkeit des Schwellwertes für die Amplitudenhöhe

erreichbare Wert für das F-Maß während des Rennens liegt bei 0,224. Im Vergleich hierzu erreicht das F-Maß während Phasen mit geringer körperlicher Anstrengung weitaus höhere Werte von 0,552 (Tippen links) bzw. 0,539 (Tippen rechts). Des Weiteren fällt auf, dass die optimale Amplitudenhöhe zur Trennung emotionaler von nicht-emotionalen Reaktionen mit zunehmender körperlicher Aktivität abnimmt. Dies lässt sich durch die im Zusammenhang mit körperlicher Aktivität insgesamt niedrigere Amplitudenhöhe erklären.

Zusammengenommen über alle Phasen ergibt sich ein optimaler Schwellwert von 0,37  $\mu\text{S}$  zur Trennung zwischen emotionalen Reaktionen und sonstigen Reaktionen. Das F-Maß liegt hierbei bei einem Wert von 0,322.

Im Rahmen dieser Studie erfolgte gezielt die Kombination emotionaler Stimuli mit verschiedenen Intensitäten der körperlichen Aktivität. Wie die vorstehenden Ergebnisse gezeigt haben, erschwert körperliche Aktivität die Detektion der emotionalen Stimuli. Zusätzlich wurde in dieser Studie die körperliche Aktivität am Trageort des EDA-Sensors sowie das

EKG-Signal aufgezeichnet. Nachfolgende Analysen sollen nun untersuchen, inwiefern die Erkennungsgenauigkeit durch die Hinzunahme von Kontextinformationen aus diesen beiden Signalen verbessert werden kann. Ziel der Analysen ist, einen möglichst großen Teil der Daten, welche körperliche Aktivität enthalten, zu verwerfen, dabei jedoch einen möglichst großen Anteil der Daten, bei welchen wenig oder nur geringe körperliche Aktivität stattfindet, zu behalten. Für die Analysen wurde jeweils eine Fensterung mit einer Fenstergröße von 15 Sekunden auf dem Aktivitäts- und dem Herzratensignal durchgeführt, um kurzfristige Störungen des Signals, wie sie beispielsweise durch Zusammenzucken bei einem Stimulus zustande kommen können, weitestgehend unberücksichtigt zu lassen.



**Abbildung 6.5:** Anteil der akzeptierten EDA-Daten in Abhängigkeit von Aktivität (oben) und Herzrate (unten)

In Abbildung 6.5 ist der Anteil der nicht verworfenen EDA-Daten in Abhängigkeit der Schwellwerte für Aktivität und Herzrate dargestellt. Die Abbildung zeigt, dass sich zwar die Phase des Rennens über die Aktivität sehr gut von den restlichen Signalen trennen lässt, jedoch ist die Ergometer-Phase – in welcher zwar starke körperliche Anstrengung vorhanden ist, jedoch nur geringe Bewegung auftritt – nur schwer von den restlichen Phasen zu unterscheiden. Des Weiteren tritt für das Karten Sortieren ein ähnlich hohes Aktivitätssignal auf wie auch für das Gehen, was auf den Trageort des Sensors am Handgelenk zurückzuführen ist. Somit ist es unumgänglich, zur besseren Differenzierung zwischen den Phasen ebenfalls die relative Veränderung der Herzrate im Vergleich zur Baseline-Phase mit hinzu zu ziehen.

Wählt man für die Aktivität einen oberen Schwellwert von 0,3 g, erhält man bereits eine Steigerung des F-Maßes um rund 8 % auf 0,347. Nimmt man zusätzlich einen Schwellwert für die maximale Herzratenveränderung von ebenfalls 0,3 hinzu, so steigt das F-Maß um weitere 15 % auf 0,401. Tabelle 6.1 zeigt den Anteil der EDA-Reaktionen, welche nach Anwendung dieser Kriterien in den jeweiligen Phasen behalten werden. Zwar steigt der F-Wert für niedrigere Aktivitäts- und Herzratenschwellwerte weiter an, dies bringt jedoch mit sich, dass selbst in Phasen mit geringer Bewegung wie der Baseline-Phase ein Großteil der Daten verworfen werden muss. Dies ist an dieser Stelle jedoch nicht erwünscht.

Grundsätzlich gilt festzuhalten, dass die Grenze für die jeweiligen Schwellwerte stark von dem jeweiligen Anwendungsfall abhängig ist. Tritt körperliche Anstrengung in einem Szenario nur in Verbindung mit körperlicher Aktivität auf, kann auf das EKG-Signal als Kontextsignal verzichtet werden. Sind jedoch Aktivitäten wie das Benutzen eines Ergometers absehbar, sollte das EKG-Signal auf jeden Fall hinzugezogen werden. Zudem muss von Fall zu Fall entschieden werden, ob es wichtiger ist, einen Großteil der Daten zu akzeptieren oder aber ob eine möglichst große Genauigkeit bei der Erkennung erforderlich ist und die Daten schon bei geringen Aktivitätsschwellwerten verworfen werden sollen. Ein derartiges Szenario wäre beispielsweise bei Laborstudien denkbar, bei welchen ohnehin nur geringe Bewegungsaktivitäten auftreten sollten.

	Aktivität $< \infty$ & Herzrate $< \infty$		Aktivität $< 0,3$ & Herzrate $< \infty$		Aktivität $< 0,3$ & Herzrate $< 0,3$	
	Daten	F-Maß	Daten	F-Maß	Daten	F-Maß
Baseline 1	100 %	0,327	100 %	0,327	88 %	0,364
Tippen links	100 %	0,449	100 %	0,449	95 %	0,435
Tippen rechts	100 %	0,512	100 %	0,512	100 %	0,512
Sortieren	100 %	0,317	78 %	0,313	72 %	0,333
Gehen	100 %	0,273	98 %	0,277	63 %	0,286
Rennen	100 %	0,164	0 %	–	0 %	–
Ergometer	100 %	0,205	99 %	0,208	23 %	0,316
Baseline 2	100 %	0,464	100 %	0,464	77 %	0,533
<i>Gesamt</i>	<i>100 %</i>	<i>0,322</i>	<i>82 %</i>	<i>0,347</i>	<i>59 %</i>	<i>0,401</i>

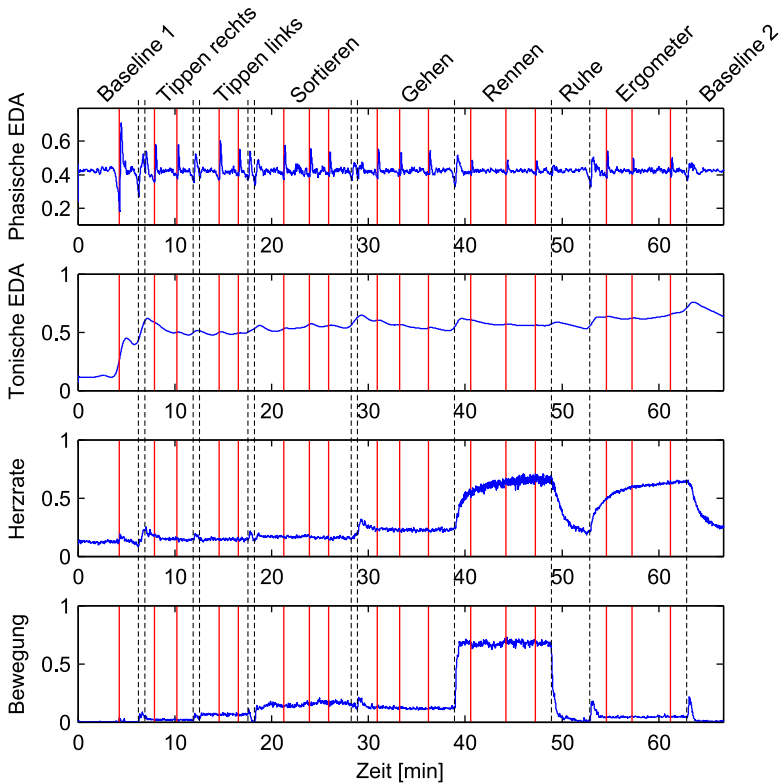
**Tabelle 6.1:** Anteil der akzeptierten Daten sowie das dazugehörige F-Maß für unterschiedliche Schwellwerte für Aktivität und Herzrate

### 6.2.3 Wirkung emotionaler Stimuli auf EDA und Herzrate

Im folgenden Abschnitt soll die grundsätzliche Wirkung emotionaler Stimuli auf den Hautleitwert sowie die Herzrate verglichen werden. Für diese Untersuchungen werden die Mittelwerte über alle Versuchspersonen gebildet. In Abbildung 6.6 ist der Mittelwert des tonischen und phasischen Anteils des EDA-Signals über alle Versuchspersonen während der Studie dargestellt. Zusätzlich wurde die mittlere Herzrate über alle Versuchspersonen ermittelt. Die emotionalen Stimuli während des Experiments sind mit roten Linien gekennzeichnet. Da bei der Mittelwertbildung absolute Werte betrachtet werden und die Wertebereiche der physiologischen Daten der einzelnen Versuchspersonen teilweise unterschiedlich ausfallen, wurde vor der Berechnung des Gruppenmittelwerts für jede Versuchsperson eine Normalisierung der physiologischen Signale auf einen Wertebereich von 0 bis 1 vorgenommen.

Die Mittelwertbildung zeigt deutlich, dass sich die einzelnen emotionalen Stimuli trotz der körperlichen Anstrengung deutlich in der phasischen elektrodermalen Aktivität widerspiegeln. Zwar nimmt die mittlere Amplitudenhöhe der Reaktionen mit höherer Anstrengung ab, ist jedoch immer noch gut zu erkennen. In der Herzrate hingegen spiegeln sich die emotionalen Stimuli nicht mit der gleichen Intensität wider. Phasenübergänge,





**Abbildung 6.6:** Normierter Mittelwert über die physiologischen Parameter aller Versuchspersonen. Rote Linien kennzeichnen emotionale Stimuli während der Durchführung des Experiments

welche mit Bewegung verbunden sind, sind jedoch in der Herzratenveränderung sehr gut zu erkennen.

Bei der tonischen Aktivität ist insbesondere der Stimulus in der Baseline-Periode gut zu erkennen. Jedoch variiert die tonische Aktivität nur recht langsam. Selbst in Phasen starker körperlicher Anstrengung wie beim Rennen oder Ergometer fahren steigt die tonische Aktivität nicht über ein bestimmtes Niveau. Jedoch schlägt sich die körperliche Anstrengung sehr

deutlich in der Herzrate nieder.

An den gemittelten Daten zeigt sich somit deutlich, dass bei einer ausreichend großen Population phasische EDA selbst bei hoher körperlicher Aktivität gut zur Identifikation kurzfristiger emotionaler Ereignisse geeignet ist, während die Herzrate sehr stark von körperlicher Aktivität überlagert wird. Auch schlagen sich einzelne Stimuli nicht im gleichen Maße nieder wie bei der phasischen EDA.

### 6.3 Zusammenfassung der Ergebnisse

Vorstehende Analysen haben gezeigt, dass das EDA-Signal grundsätzlich zur Identifikation emotionaler Stimuli geeignet ist. Jedoch nimmt die Eignung mit zunehmender körperlicher Belastung deutlich ab. Die Hinzunahme von Kontextinformationen wie der Aktivität oder der Herzratenveränderung kann jedoch helfen, die Genauigkeit der Identifikation emotionaler Stimuli zu verbessern. Der Schwellwert für das Verwerfen artefaktbehafteter Signale muss jedoch sorgfältig in Abhängigkeit der individuellen Anforderungen der Anwendung gewählt werden. Für Lernanwendungen kann die Analyse des EDA-Signals somit genutzt werden, um Zeitpunkte plötzlicher Unterbrechungen zu identifizieren und entsprechende Maßnahmen wie eine Unterbrechung des Lernprogramms einzuleiten.

Zudem hat sich gezeigt, dass bei einer Mittelwertbildung über die synchron aufgezeichneten Signale mehrerer Personen phasische EDA-Aktivität als guter Indikator für kurzfristige emotionale Stimuli geeignet ist. In der relativen Veränderung der Herzrate hingegen spiegeln sich diese kaum wider. Ein mögliches Einsatzszenario stellt das Lernen in Gruppen dar. Durch die Mittelwertbildung über das Signal mehrerer Personen sind emotionale Reaktionen im phasischen EDA-Signal selbst in Phasen hoher körperlicher Belastung deutlich erkennbar.



# 7 Nutzung von physiologischem Feedback zum Emotionsregulationstraining

Wie in Abschnitt 2.2 dargestellt, beeinflussen Emotionen viele Bereiche des täglichen Lebens. Die Fähigkeit, die eigenen Emotionen bewusst in die richtige Richtung zu steuern, ist daher von großer Bedeutung. Bevor eine Verbesserung der Emotionsregulationsfähigkeiten überhaupt stattfinden kann, ist es wichtig, ein Bewusstsein für den aktuellen emotionalen Zustand zu erlangen. Hierbei kann physiologisches Feedback helfen, da dieses ein objektives Bild des aktuellen Nutzerzustands geben kann. Neben dem Bewusstwerden des aktuellen emotionalen Zustands kann es auch dabei unterstützen, ein Gefühl für die bewusste Regulation von Emotionen zu erlangen.

Im Folgenden wird eine im Rahmen des Projektes xDelia durchgeführte Studie vorgestellt, bei welcher untersucht wird, inwiefern physiologisches Feedback dazu beitragen kann, das persönliche emotionale Erregungsniveau effizienter zu regulieren. In Abschnitt 7.1 wird zunächst der Studienaufbau sowie die Durchführung der Studie beschrieben. Anschließend folgt in Abschnitt 7.2 die Analyse der Daten. Das Kapitel schließt mit einer Zusammenfassung der wichtigsten Ergebnisse dieser Studie in Abschnitt 7.3.

Teile dieses Kapitels wurden bereits in abgeänderter Form in Jerčić et al. (2012); Astor et al. (2014) veröffentlicht.

## 7.1 Studiendesign

Nachfolgend ist der genaue Aufbau und Ablauf der Studie beschrieben. In Abschnitt 7.1.1 sind zunächst die Konzeption der Studie sowie die

verwendeten Hard- und Softwarelösungen dargestellt, während in Abschnitt 7.1.2 auf den Ablauf der Studie eingegangen wird.

### 7.1.1 Studienaufbau

Ziel dieser Studie ist es zu untersuchen, inwiefern physiologisches Feedback sinnvoll zum Training der Emotionsregulationsfähigkeiten beitragen kann, da dieses helfen kann, ein besseres Bewusstsein für den aktuellen emotionalen Zustand zu erhalten. Zudem soll untersucht werden, ob sich die Trainingseffekte auch auf weitere Transfer-Aufgaben auswirken.

Da die Fähigkeit zur effektiven Emotionsregulation insbesondere bei Handelsentscheidungen an Finanzmärkten eine essenzielle Rolle spielt, wurde als Kern der Studie ein Spiel verwendet, bei welchem Finanzentscheidungen an einem Aktienmarkt simuliert werden. Der Spieler nimmt hierbei die Rolle eines Aktionärs ein, der Aktien kauft und verkauft mit dem Ziel, seinen Gewinn zu maximieren. Zur Untersuchung, inwiefern sich die Trainingsergebnisse durch das physiologische Feedback verändern, wurde das Spiel in zwei Versionen eingesetzt: einer Version mit und einer ohne physiologisches Feedback. Grundgedanke bei der Feedback-Version ist, dass das Spiel mit zunehmendem Erregungsniveau des Spielers an Schwierigkeit zunimmt und somit die Fähigkeit, die emotionale Erregung auf einem niedrigeren Niveau zu halten, durch ein niedrigeres Schwierigkeitsniveau belohnt wird.

Zur Untersuchung von Transfer-Effekten wurde zum einen untersucht, ob Unterschiede bei einer ähnlichen Aufgabe zu beobachten sind, sowie ob sich die Reaktionen bei Aufgaben mit emotionalen Inhalten verändern.

In Abschnitt 7.1.1.1 wird die Software beschrieben, welche bei der Studie zum Einsatz kam. Des Weiteren wird in Abschnitt 7.1.1.2 die im Rahmen der Studie verwendete Hardware dargestellt.

#### 7.1.1.1 Software

Zur Untersuchung des Einflusses von physiologischem Feedback auf das Training des Emotionsregulationsverhaltens wurde das Auction Game in Kombination mit der xAffect-Software verwendet.

Das Auction Game wurde im Rahmen des EU-Projektes xDelia als Werkzeug entwickelt, welches Akteuren an Finanzmärkten helfen soll, ihre

Emotionsregulationsfähigkeiten zu trainieren (Jerčić et al., 2012). Das Spiel bildet ein finanzmarktähnliches Szenario nach.

Bei der Entwicklung des Spiels standen folgende Anforderungen im Vordergrund (Astor et al., 2014):

- Das Spiel soll emotionale Erregung auslösen und ist in einen Finanzkontext eingebunden
- Effiziente Emotionsregulation soll belohnt werden
- Das Spiel soll dem Nutzer möglichst belastungsfrei aussagekräftiges Feedback über dessen Zustand geben.

Im Rahmen einer weitreichenden Evaluation dieses Spiels (Astor et al., 2014) konnte gezeigt werden, dass das Spiel allen drei Anforderungen genügt. So konnte während des Spiels eine signifikante Steigerung der Herzrate um durchschnittlich 11,07% festgestellt werden. Auch zeigte sich ein signifikanter negativer Einfluss des Erregungsniveaus auf die Entscheidungsperformance. Zudem war das mittlere Erregungsniveau der Versuchspersonen, welchen im Spiel Feedback über ihr emotionales Erregungsniveau gegeben wurde, niedriger als das der Versuchspersonen, welche das gleiche Spiel ohne Feedback spielten.

Das Grundprinzip des Spiels besteht darin, dass der Spieler in jeder Runde drei Preisabschätzungen für eine Aktie präsentiert bekommt. In Abbildung 7.1 ist dies beispielhaft dargestellt. Der wahre Preis kann als Mittelwert aus diesen drei Werten berechnet werden. Basierend hierauf muss der Spieler die Entscheidung treffen, ob er eine Aktie zu einem vorgegebenen Preis kaufen oder verkaufen möchte. Zusätzlich steht dem Spieler über den *No-Trade-Button* die Möglichkeit zur Verfügung, in einer Runde keine Entscheidung zu treffen. Für jede Handelsentscheidung steht nur eine begrenzte Zeit zur Verfügung, welche mittels der Uhr am linken oberen Rand des Bildschirms angezeigt wird. Trifft ein Spieler innerhalb dieser Zeit keine Entscheidung und betätigt nicht den *No-Trade-Button*, so führt dies zu dem maximal möglichen Verlust von 5 Euro. Die Aufgabe des Spielenden besteht nun darin, im Verlauf des Spieles möglichst viel Geld zu verdienen, was durch eine hohe Anzahl an richtigen Entscheidungen erreicht werden kann. Insgesamt müssen 7 Level durchlaufen werden. Werden in einem Level die zeitlichen oder monetären Ziele nicht erfüllt, endet das Spiel vorzeitig.



Abbildung 7.1: Spielaufbau des Auction Games (Astor et al., 2014)

Das Spiel steht in zwei verschiedenen Varianten zur Verfügung: einer Variante mit und einer ohne physiologischem Feedback. In der Variante ohne physiologisches Feedback werden zwar physiologische Signale für spätere Analyse Zwecke aufgezeichnet, jedoch haben diese keinen Einfluss auf den Spielverlauf. Auch erfolgt keine Anzeige des aktuellen Erregungsniveaus des Spielers. In der Version mit physiologischem Feedback hingegen wird die Schwierigkeit des Spiels in Abhängigkeit des emotionalen Erregungsniveaus eines Spielers variiert: Je höher das Erregungsniveau eines Spielers, desto schwieriger wird auch das Spiel. Erfolgreiche Emotionsregulation wird somit durch eine niedrigere Spielschwierigkeit belohnt. Das aktuelle Erregungsniveau wird dem Spieler auf einer Erregungsskala angezeigt. Diese zeigt die verschiedenen Erregungsstufen auf einer Farbskala von grün (1) bis rot (5) an.

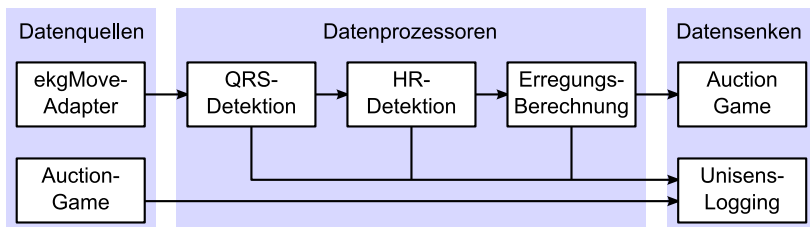
Das Erregungsniveau eines Spielers wird basierend auf der Herzrate ermittelt. Hierfür wird die Komponente zur Erregung-Berechnung aus xAffect verwendet. Vor dem eigentlichen Beginn des Auction Games erfolgt eine fünfminütige Baseline-Periode, während welcher die individuelle Baseline-

Herzrate für eine Versuchsperson ermittelt wird. Das Erregungsniveau wird anschließend aus der relativen Veränderung der Herzrate während des Spiels zur Baseline-Herzrate abgeleitet. Um zu vermeiden, dass sich die respiratorische Sinusarrhythmie in den berechneten Erregungswerten widerspiegelt, wird zur Bestimmung des Erregungsniveaus während des Spiels jeweils ein Mittelwert über die letzten fünf Herzschläge verwendet. Das so berechnete Erregungsniveau wird mit einer Samplerate von 1 Hz von der xAffect-Software an das Auction Game gesendet. Im Spiel wird das Erregungsniveau in 5 verschiedene Spielschwierigkeiten umgesetzt, wobei 1 das leichteste und 5 das schwerste Schwierigkeitsniveau darstellt.

Folgende Spielelemente verändern sich mit zunehmendem Erregungsniveau eines Spielers:

- Die *Varianz der Preisabschätzungen* nimmt zu. Dies erschwert die Berechnung des tatsächlichen Preises einer Aktie.
- Der *Preis der Aktie in der nächsten Runde* variiert stärker von dem Preis der Vorrunde, wodurch die Preisänderungen weniger vorhersehbar sind.
- In höheren Leveln wird die *Geschwindigkeit der Wolken* erhöht. Die schnellere Bewegung der Wolken erschwert die visuelle Erfassung der Preisabschätzungen.

Für weiterführende Informationen über das Spieldesign und dessen Evaluation sei auf Astor et al. (2014) verwiesen.



**Abbildung 7.2:** Setup der xAffect-Software für das Auction Game

Abbildung 7.2 illustriert das in diesem Rahmen erstellte xAffect-Setup. Das Spiel ist hierbei sowohl als Datenquelle mit dem Zweck der Aufzeich-



nung der relevanten Spieldaten eingebunden, als auch als Datensenke, da über das Netzwerk Informationen von xAffect an das Spiel gesendet werden. Des Weiteren werden während des Spiels die physiologischen Signale eines Spielers sowie die daraus abgeleiteten Parameter aufgezeichnet .

Des Weiteren wurde zur Untersuchung von Transfer-Effekten auf ähnliche Aufgaben das Auctio Game Diagnostic verwendet. Das Spielprinzip ist ähnlich aufgebaut wie bei der didaktischen Version. Jedoch spielen die Spieler jeweils 10 Runden des Spiels welche aus je 5 Entscheidungen bestehen. Nach jeder Entscheidung erhält der Spieler Feedback, ob seine Entscheidung richtig oder falsch war sowie über den Gewinn oder Verlust, der durch diese Entscheidung entstanden ist. Zusätzlich wird ein zusammenfassendes Feedback nach jeder Runde des Spiels angezeigt. Wie auch beim eigentlichen Auction Game erfolgt bei der diagnostischen Variante eine Aufzeichnung der physiologischen Signale des Spielers. Jedoch hat dieses Signal keinen Einfluss auf die Schwierigkeit des Spiels.

Für die Untersuchung der Veränderungen bei der Reaktion auf Aufgaben mit emotionalen Inhalten kam die Software E-Prime zum Einsatz. Diese Software stellt eine Entwicklungsumgebung für die Durchführung von computerbasierten verhaltenspsychologischen Experimenten zur Verfügung und wurde im Rahmen der Studie zur Durchführung des emotional Stroop-Tests verwendet.

### 7.1.1.2 Hardware

Im Rahmen dieser Studie war die Hauptanforderung an die Sensoren, dass diese möglichst belastungsfrei über eine längere Zeit getragen werden können, da das Spiel später auch außerhalb von Laborstudien eingesetzt werden soll. Zudem ist eine gute Signalqualität von großer Bedeutung, damit das emotionale Erregungsniveau möglichst genau bestimmt werden kann. Aus diesem Grund wurde der in Abschnitt 4.2.1 beschriebene ekgMove-Sensor verwendet, da dieser die gestellten Anforderungen am besten erfüllt.

### 7.1.2 Studienablauf

Die Studie wurde in drei Sessions mit Abstand von jeweils einer Woche am KIT durchgeführt. Insgesamt nahmen 107 Versuchspersonen (davon 27

weiblich) an der Studie teil. Das Durchschnittsalter der Versuchspersonen lag bei 22,14 (SD = 2,18) Jahren. Sämtliche Versuchspersonen hatten keine Erfahrungen mit Emotionsregulationstraining. Für die Teilnahme an allen drei Sessions erhielten die Versuchspersonen eine Vergütung von 40 Euro. Neben der fixen Vergütung gab es eine variable Vergütung zwischen 0 und 30 Euro in Abhängigkeit von der Performance während der einzelnen Sessions. Die Einteilung der Versuchspersonen erfolgte zufällig in die nachfolgenden Gruppen:

**CNT** Die Kontrollgruppe spielt das Auction Game, ohne dass die physiologische Erregung einen Einfluss auf das Spiel ausübt. Das Erregungsniveau wird nicht angezeigt

**AUC** Diese Gruppe spielt ausschließlich das Auction Game mit physiologischem Feedback

**ASI** Diese Gruppe spielt zunächst das Auction Game mit physiologischem Feedback sowie anschließend das Spiel Space Investors, welches ebenfalls zur Emotionsregulation entwickelt wurde

**SIA** In dieser Gruppe wird Space Investors vor dem Auction Game mit physiologischem Feedback gespielt

In jeder Woche nahmen die Versuchspersonen an der ihrer Gruppeneinteilung entsprechenden Intervention teil. Zur Erfassung des Emotionsregulationsverhaltens wurde der Emotion Regulation Questionnaire (ERQ) von Gross & John (2003) verwendet. Im Anschluss an die jeweilige Intervention wurden mittels des State-ERQ die während des Spiels verwendeten Emotionsregulationsstrategien erfasst. Hierbei handelt es sich um eine leicht modifizierte Variante des ERQ, welche explizit auf das Spiel bezogen war. Im Anschluss an den Fragebogen folgten zwei weitere Aufgaben zur Untersuchung von Transfer-Effekten, auf welche in Abschnitt 7.2.5 genauer eingegangen wird. Der Ablauf einer Session ist in Abbildung 7.3 graphisch veranschaulicht.

Eine Woche vor Beginn der Studie sowie drei Wochen nach der Studie erhielten alle Teilnehmer den Trait-ERQ, um zusätzlich zu den Informationen über Emotionsregulation während des Spiels Informationen über die Emotionsregulationseigenschaften der Personen zu erhalten.

<b>CNT</b>	Auction Game (ohne Biofeedback)		Emotion Regulation Questionnaire	Emotional Stroop Test	Auction Game Diagnostic
<b>AUC</b>	Auction Game				
<b>ASI</b>	Auction Game	Space Investors			
<b>SIA</b>	Space Investors	Auction Game			

**Abbildung 7.3:** Ablauf einer einzelnen Session

## 7.2 Datenanalyse

Im Folgenden sind die Ergebnisse der Datenanalysen dargestellt. Bei den Analysen wurden nur diejenigen Daten berücksichtigt, bei welchen für alle drei Wochen des Experiments physiologische Daten vorlagen. Alle anderen Datensätze wurden von der Analyse ausgeschlossen. Da einzelne Teilnehmer nicht zu allen drei Sitzungen erschienen, sowie aufgrund technischer Probleme während der Aufnahme der EKG-Daten, standen für die nachfolgenden Analysen insgesamt 86 vollständige Datensätze zur Verfügung, wovon 23 Datensätze auf die Kontrollgruppe sowie 63 Datensätze auf die ER-Gruppe (genauer ASI: 21, AUC: 20, SIA: 22) entfallen. 22 dieser Versuchspersonen sind weiblich, der Rest männlich. Das Durchschnittsalter liegt bei 22,22 (SD = 2,23).

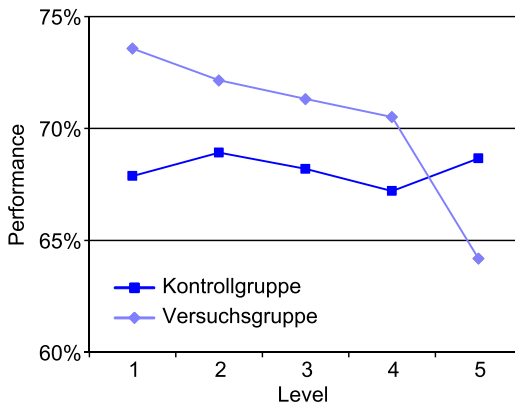
Die Spielperformance  $p$  eines Spielers errechnet sich wie folgt aus der Anzahl der richtigen und falschen Entscheidungen sowie den Timeouts:

$$p = \frac{N_{\text{correct}}}{N_{\text{correct}} + N_{\text{wrong}} + N_{\text{timeout}}} \quad (7.1)$$

Das Betätigen des Notrade-Buttons wird hierbei nicht mit in Betracht gezogen.

### 7.2.1 Einfluss der Spielschwierigkeit auf die Performance

Wie einführend erläutert, kann das Auction Game während des Spielverlaufs fünf verschiedene Schwierigkeitsstufen annehmen, die sich basierend auf der relativen Abweichung der Herzrate von der mittleren Herzrate während der Baseline berechnen. Der Schwellwert für die Berechnung der Spielschwierigkeit wurde bei dem Entwurf des Spiels so festgelegt, dass die höchste Schwierigkeitsstufe nach Möglichkeit von jeder Versuchsperson erreicht wird. Wie die Auswertungen der ER-Gruppe zeigen, erreichten in der 1. Woche noch 96,00 % aller Versuchspersonen die höchste Schwierigkeitsstufe. Aufgrund von Trainingseffekten wurde Stufe 5 in der 2. Woche nur noch von 83,12 % und in der 3. Woche von 87,67 % aller Versuchspersonen erreicht.



**Abbildung 7.4:** Zusammenhang von Performance und Spielschwierigkeit für beide Gruppen

Abbildung 7.4 zeigt den Anteil der korrekt getroffenen Entscheidungen an den Gesamtentscheidungen in Abhängigkeit von der Schwierigkeitsstufe des Spiels für alle Spieler, aufgeteilt nach Versuchsgruppe und Kontrollgruppe. Wie erwartet, nimmt die Performance der Spieler in der ER-Gruppe mit steigender Spielschwierigkeit deutlich ab. Bei der Kontrollgruppe hingegen sind nur sehr geringe Unterschiede zu beobachten, da hier die theoretische Spielschwierigkeit basierend auf dem Erregungsniveau zwar berechnet wird,

jedoch findet keine tatsächliche Veränderung der Spielschwierigkeit statt.

### 7.2.2 Einsatz von Emotionsregulation während des Spiels

Ziel dieser Studie war, bei den Versuchspersonen eine Veränderung in ihrem Emotionsregulationsverhalten hervorzurufen. Es soll daher überprüft werden, inwiefern sich Versuchsgruppe und Kontrollgruppe in ihren Emotionsregulationsstrategien während des Spiels unterscheiden. Die erste in diesem Rahmen untersuchte Fragestellung ist, inwiefern das wiederholte Spielen des Auction Games über die Wochen hinweg zu einer Veränderung des Emotionsregulationsverhaltens der Versuchspersonen führt. Des Weiteren wird angenommen, dass das Spielen des Auction Games mit physiologischem Feedback bei den Versuchspersonen der AUC-Bedingung dazu führt, dass gegenüber der Kontrollgruppe während des Spiels mehr Emotionsregulationsstrategien angewendet werden. Beide Fragestellungen sollen anhand der physiologischen Daten, welche während des Spiels sowie der vorhergehenden Baseline aufgezeichnet wurden, untersucht werden. Des Weiteren werden auch die Fragebogendaten herangezogen, welche im Anschluss an das Spiel gesammelt wurden.

Zur Untersuchung dieser Fragestellung werden ausschließlich die Daten aus der CNT- und der AUC-Intervention herangezogen, da nicht ausgeschlossen werden kann, dass durch das Spielen von Space Investors zusätzliche Wechselwirkungen mit den physiologischen Parametern sowie den Ergebnissen des Fragebogens erzeugt werden. So könnte das vorhergehende Spielen von Space Investors (SIA-Bedingung) dazu führen, dass die Versuchspersonen direkt mit einem höheren Erregungsniveau in das Auction Game starten. Wird Space Investors anschließend gespielt (ASI-Bedingung), ist ein Einfluss dieses Spiels auf den ERQ nicht auszuschließen. Diese Wechselwirkungen sind jedoch nicht Bestandteil dieser Untersuchungen. Zudem wurden nur Daten derjenigen Versuchspersonen verwendet, für welche aus allen drei Wochen vollständige Fragebogen- und physiologische Daten zur Verfügung standen, wodurch sich eine Stichprobengröße von  $N_{AUC} = 19$  und  $N_{CNT} = 23$  ergibt.

Die Anwendung der beiden Emotionsregulationsstrategien Reappraisal und Suppression soll mittels der Ergebnisse des State-ERQ untersucht werden. Zusätzlich zu den Ergebnissen der Fragebögen werden in den Analysen die relative mittlere Veränderung der Herzrate einer Person

während des Spiels im Verhältnis zur Baseline herangezogen sowie die HRV, da diese – wie bereits in Abschnitt 3.2.3.1 erläutert – Rückschlüsse auf die Emotionsregulationsfähigkeiten einer Person erlaubt. Erfolgreiche Emotionsregulation drückt sich üblicherweise in einer hohen HRV aus. Es wird zum einen die HRV während der Baseline als auch die mittlere HRV für jede Versuchsperson über die einzelnen Level des Spiels betrachtet. Da die durchschnittliche Dauer eines Levels bei rund 120 Sekunden liegt, wird als Maß für die HRV der Parameter RMSSD herangezogen, da HRV-Parameter im Frequenzbereich für Analysen von kurzen Zeitfenstern unter 5 Minuten nur bedingt geeignet sind (siehe hierzu auch Abschnitt 5.2.2.2). Ähnlich wie der Parameter HF reflektiert auch RMSSD den hochfrequenten Anteil der HRV (Goedhart et al., 2007), ist jedoch auch für kürzere Analysezeitfenster, wie sie während der einzelnen Level des Spiels auftreten, stabil. Neben der HRV während des Spiels werden auch Unterschiede der HRV zwischen beiden Gruppen während der Baseline-Periode untersucht.

		Woche 1	Woche 2	Woche 3
HR <sub>diff</sub>	CNT	0,15 (0,11)	0,12 (0,11)	0,08 (0,09)
	AUC	0,07 (0,09)	0,04 (0,06)	0,02 (0,06)
RMSSD <sub>base</sub>	CNT	44,02 (13,94)	42,13 (21,84)	40,60 (23,78)
	AUC	59,27 (28,17)	49,26 (30,88)	48,26 (23,78)
RMSSD <sub>game</sub>	CNT	29,09 (12,40)	30,19 (18,47)	33,83 (16,97)
	AUC	47,63 (29,42)	43,76 (31,43)	45,72 (29,06)
Reappraisal	CNT	3,75 (1,12)	3,69 (1,76)	3,38 (1,29)
	AUC	3,95 (1,22)	4,26 (1,37)	4,11 (1,35)
Suppression	CNT	3,77 (0,82)	3,51 (1,11)	3,85 (1,28)
	AUC	4,61 (1,33)	4,49 (1,33)	4,52 (1,30)

**Tabelle 7.1:** Mittelwerte und Standardabweichung für die abhängigen Variablen

In Tabelle 7.1 sind die Mittelwerte der abhängigen Variablen in Abhängigkeit der Versuchsbedingung sowie der Zeit dargestellt. Es zeigt sich, dass die mittlere Herzratenveränderung im Spiel gegenüber der Baseline für die Kontrollgruppe weitaus höher ausfällt, als für die Versuchspersonen in der AUC-Bedingung. Für beide Gruppen sinkt diese jedoch über die Wochen. Die RMSSD-Werte reflektieren, dass die Versuchspersonen aus der AUC-Bedingung sowohl während der Baseline, als auch während des Spiels höhere HRV-Werte aufweisen. Jedoch ist über die Wochen hinweg ein leicht-

tes Absinken von RMSSD zu beobachten. Lediglich in der Kontrollgruppe kommt es während des Spiels zu einem leichten Anstieg von RMSSD. Für den ERQ kann bezüglich des Reappraisal-Wertes bei der Kontrollgruppe im Verlauf der Zeit ein Absinken beobachtet werden, während der Wert bei der AUC-Bedingung ein leichter Anstieg erfolgt. Der Suppression-Wert für die AUC-Bedingung hingegen sinkt im Verlauf über die Zeit leicht ab, wohingegen die Kontrollgruppe am Ende der Studie einen höheren Suppression-Wert aufweist als zu Beginn.

Zur Überprüfung, inwiefern es Veränderungen des Emotionsregulationsverhaltens über die Zeit gibt, sowie zur Analyse der Unterschiede zwischen den beiden Gruppen, wird eine MANOVA mit Messwiederholung mit dem unabhängigen Faktor der Versuchsbedingung durchgeführt. Als abhängige Variablen werden die mittlere Herzratenveränderung während des Spiels im Vergleich zur Baseline, RMSSD während der Baseline und im Mittel während der einzelnen Level des Spiels sowie die Ergebnisse des ERQ für Suppression und Reappraisal während der drei Wochen herangezogen.

Die MANOVA zeigt einen signifikanten multivariaten Haupteffekt für den Faktor Zeit ( $F(10; 31) = 2,574, p < 0,05$ ). Jedoch können keine signifikanten Interaktionseffekte zwischen der Zeit und der Versuchsbedingung festgestellt werden ( $F(10; 31) = 0,592, p = 0,808$ ), was bedeutet, dass die Gruppenunterschiede über die Zeit hinweg stabil bleiben.

Faktor	df	F	Sig
HR <sub>diff</sub>	1,724	8,420	0,001***
RMSSD <sub>base</sub>	2	2,530	0,086*
RMSSD <sub>game</sub>	2	0,589	0,557
Reappraisal	2	0,455	0,636
Suppression	2	0,402	0,670

\*  $p < 0,1$ ; \*\*  $p < 0,05$ ; \*\*\*  $p < 0,01$

**Tabelle 7.2:** Univariate Tests für die abhängigen Variablen

Vor dem Hintergrund des signifikanten Haupteffektes werden die abhängigen Variablen in detaillierten ANOVAS weiter untersucht. Die Ergebnisse sind in Tabelle 7.2 dargestellt. Aufgrund einer Verletzung der Sphärizität für die HR-Veränderung wurde hierfür eine Greenhouse-Geisser-Korrektur vorgenommen. Es zeigt sich, dass der Faktor Zeit insbesondere auf die mittlere Herzrate während des Spiels einen signifikanten Einfluss ausübt.

Auch bei der HRV während der Baseline zeigt sich ein leichter Einfluss. Alle anderen Faktoren weisen über den Verlauf der Wochen keine signifikanten Veränderungen auf. Es kann somit nicht bestätigt werden, dass das wiederholte Spielen des Spiels eine Veränderung des Emotionsregulationsverhaltens hervorruft, da sich dies auch in den Ergebnissen des State-ERQ sowie der HRV während des Spiels hätte ausdrücken müssen. Die signifikante Veränderung der HR allein reicht in diesem Fall nicht als Indiz für ein verändertes Emotionsregulationsverhalten aus, da diese auch allein auf Trainingseffekten beruhen kann.

Bezüglich der Unterschiede zwischen den beiden Versuchsgruppen zeigt der Test der Zwischenssubjekteffekte für den Faktor Versuchsbedingung signifikante Effekte für die HR-Veränderung ( $F(1; 40) = 8,593$ ,  $p < 0,01$ ), RMSSD während des Spiels ( $F(1; 40) = 2,131$ ,  $p < 0,05$ ) sowie Suppression ( $F(1; 40) = 12,482$ ,  $p < 0,01$ ). Für RMSSD während der Baseline ( $F(1; 40) = 2,131$ ,  $p = 0,152$ ) und den Reappraisal-Wert ( $F(1; 40) = 2,432$ ,  $p = 0,127$ ) sind zwar Unterschiede erkennbar, jedoch sind diese nicht signifikant. Dies lässt darauf schließen, dass das Spielen des Spiels mit physiologischem Feedback dazu führt, dass signifikant häufiger die Emotionsregulationsstrategie Suppression angewendet wird, was sich auch in der HRV während des Spiels widerspiegelt, welche für Versuchspersonen der AUC-Bedingung höher ist als in der Kontrollgruppe. Während der Baseline reflektieren sich diese Unterschiede bzgl. der Emotionsregulation jedoch nicht in der HRV. Auch sind die Unterschiede für die Strategie Reappraisal zwischen den Gruppen nicht signifikant. Die stärkere Anwendung von Suppression kann als möglicher Grund dafür gesehen werden, dass auch die mittlere Herzratenveränderung zwischen den beiden Versuchsgruppen für die AUC-Bedingung geringer ausfällt als für die Kontrollgruppe.

Somit kann die Hypothese, dass in der Versuchsgruppe mehr Emotionsregulationsstrategien angewendet werden als in der Kontrollgruppe, nur teilweise bestätigt werden, da die Unterschiede nur bei der Strategie Suppression nicht jedoch bei Reappraisal zu beobachten sind. Auch die Herzrate sowie RMSSD während des Spiels spiegeln den unterschiedlichen Gebrauch von Emotionsregulationsstrategien wider.



### 7.2.3 Einfluss des ER-Trainings auf die Spielperformance

Ziel des Spiels ist es, die Emotionsregulationsfähigkeiten einer Person zu verbessern, damit diese in einer Handlungsumgebung bessere Entscheidungen treffen kann. Wie bereits bei Astor et al. (2014) gezeigt, steigt die Performance der Versuchspersonen beim Spielen des Auction Games bei abnehmendem Erregungsniveau. Da sich das Erregungsniveau aus der Herzrate ableitet, wird nachfolgend der Zusammenhang zwischen der mittleren Herzratenveränderung während des Spiels zur Spielperformance untersucht. Hierzu wird eine lineare Regression durchgeführt. Es werden alle Versuchspersonen aus den Gruppen AUC und CNT berücksichtigt, für welche in allen drei Wochen physiologische Daten aufgezeichnet wurden. Hieraus ergibt sich eine Stichprobengröße von  $N_{CNT} = 23$  und  $N_{AUC} = 20$ . Die Ergebnisse der Regression sind in Tabelle 7.3 dargestellt.

Treatment	Messzeitpunkt	B	SE	T	Sig
AUC	Woche 1	-0,869	0,252	-3,454	0,003**
	Woche 2	-0,757	0,398	-1,902	0,073*
	Woche 3	-0,647	0,447	-1,446	0,165
CNT	Woche 1	-0,376	0,252	-1,490	0,151
	Woche 2	-0,435	0,222	-1,959	0,064*
	Woche 3	-0,457	0,141	-3,250	0,004**

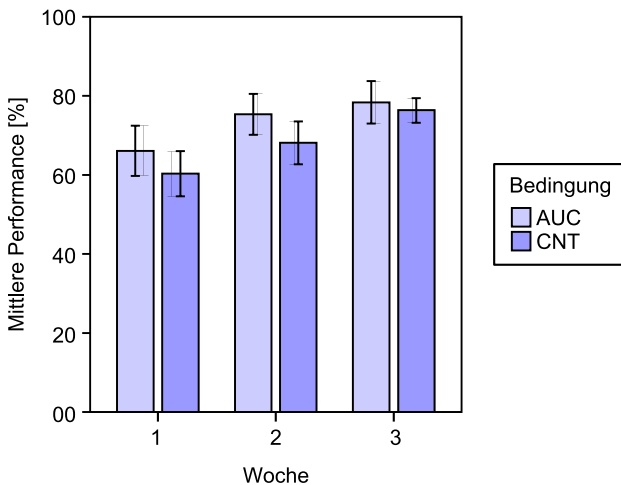
\*  $p < 0,1$ ; \*\*  $p < 0,05$

**Tabelle 7.3:** Regression über die Abhängigkeit der Performance von der Herzrate der Versuchspersonen

Wie bei Astor et al. (2014) zeigt sich auch hier ein negativer Zusammenhang zwischen der physiologischen Erregung gemessen an der durchschnittlichen Herzratenveränderung während des Spiels und der Spielperformance. Dieser fällt für Versuchspersonen aus der AUC-Bedingung weitaus größer aus, als für die Kontrollgruppe. Eine mögliche Erklärung für den größeren Einfluss der Herzrate auf die Performance bei der Versuchsgruppe kann darin gesehen werden, dass die Spielschwierigkeit bei der Versuchsgruppe bei steigender Herzrate zunimmt, bei der Kontrollgruppe jedoch gleich bleibt. Zudem ist zu beobachten, dass der Einfluss der Herzrate auf die Spielperformance bei der Versuchsgruppe im Verlauf der Wochen abnimmt, während dieser bei der Kontrollgruppe über die Wochen zunimmt.

Im Zuge dessen soll daher untersucht werden, ob der Performance-Zuwachs für Versuchspersonen in der AUC-Bedingung größer ist als für Versuchspersonen in der CNT-Bedingung.

Abbildung 7.5 zeigt, dass die Performance für beide Gruppen im Verlauf der Wochen steigt. Zudem fällt die durchschnittliche Spielperformance in der AUC-Gruppe durchgehend leicht besser aus, als in der Kontrollgruppe. Jedoch liegt die mittlere Performance-Steigerung von Woche 1 auf Woche 3 für die AUC-Gruppe bei 18,58 % ( $M = 0,123$ ;  $SD = 0,169$ ) wohingegen bei der Kontrollgruppe eine Performance-Steigerung von 26,53 % ( $M = 0,160$ ;  $SD = 0,133$ ) auftritt. Die Annahme, dass Versuchspersonen in der AUC-Bedingung im Verlauf der Wochen einen größeren Zuwachs in der Spielperformance aufweisen, kann somit nicht bestätigt werden. Eine mögliche Erklärung hierfür könnte darin gesehen werden, dass die bewusste Emotionsregulation zu viele kognitive Ressourcen bindet und somit weniger Kapazitäten für das Spiel verbleiben.



**Abbildung 7.5:** Durchschnittliche Spielperformance für beide Gruppen im Verlauf der Wochen (Fehlerbalken geben die Standardabweichung an)

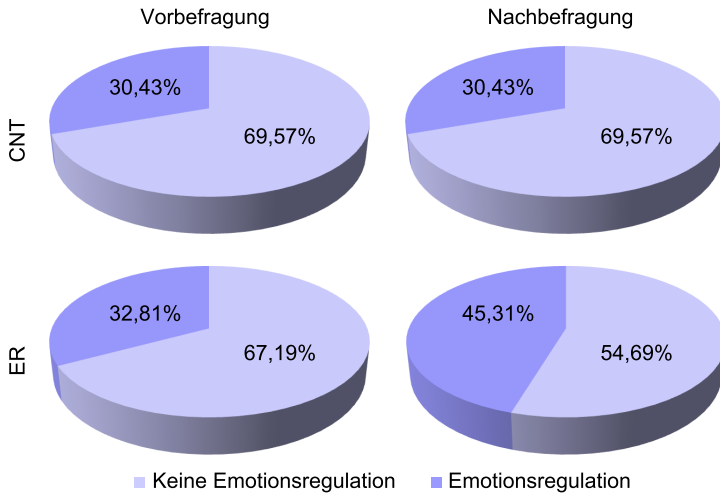
### 7.2.4 Langfristige Veränderungen des Emotionsregulationsverhaltens

Durch das Spielen des Spiels mit physiologischem Feedback soll bewirkt werden, dass sich das Emotionsregulationsverhalten der Versuchspersonen nachhaltig ändert. Es wird daher vermutet, dass das Spielen der Spiele mit physiologischem Feedback dazu führt, dass von der Versuchsgruppe auch im Nachgang zur Studie mehr Emotionsregulationsstrategien angewendet werden, als von der Kontrollgruppe.

Zur Analyse werden die Daten des Trait-ERQ herangezogen, welcher den Versuchspersonen eine Woche vor (Messzeitpunkt 1) sowie drei Wochen nach (Messzeitpunkt 2) der Durchführung der Studie vorgelegt wurde. Im Gegensatz zum State-ERQ bezieht sich dieser nicht auf die konkrete Spielsituation, sondern auf das allgemeine Emotionsregulationsverhalten. Bei der Analyse werden alle Versuchspersonen, welche einen Wert von mindestens 4 in einer der Emotionsregulationskategorien Reappraisal oder Suppression aufweisen, der Gruppe der Emotionsregulatoren zugeordnet. Alle anderen Versuchspersonen werden als Nichtregulatoren eingeordnet.

Da für diese Fragestellung im Gegensatz zu den vorherigen Analysen kurzfristige Wechselwirkungen zwischen Space Investors und dem Auction Game nicht relevant sind, werden bei den Untersuchungen die Daten aus der Kontrollgruppe sowie der ER-Gruppe (welche alle Versuchspersonen der Bedingungen AUC, ASI und SIA enthält) betrachtet. Erneut werden nur die Datensätze berücksichtigt, bei welchen in allen drei Trainingswochen physiologische Daten vorliegen und somit das Spiel mit physiologischem Feedback gespielt wurde. Für die nachfolgenden Analysen ergibt sich somit eine Stichprobengröße von  $N_{CNT} = 23$  und  $N_{ER} = 64$ .

Wie aus Abbildung 7.6 ersichtlich wird, gehören vor Durchführung der Studie etwa ein Drittel aller Versuchspersonen aus beiden Versuchsbedingungen zur Gruppe der Emotionsregulatoren. Während die Anzahl der Emotionsregulatoren in der Kontrollgruppe drei Wochen nach Abschluss der Intervention weiterhin konstant bei 30,43 % liegt, ist bei den Versuchspersonen aus der ER-Gruppe ein Anstieg der Emotionsregulatoren von 32,91 % auf 45,31 % zu beobachten. Zur Überprüfung, ob das Auftreten von Emotionsregulatoren in der Nachbefragung im Vergleich zur Vorbefragung signifikant ist, wird ein einseitiger  $z$ -Test herangezogen. Da die



**Abbildung 7.6:** Anwendung von Emotionsregulation in Versuchs- und Kontrollgruppe vor und nach der Studie

Laplace-Bedingung  $\sigma = \sqrt{64 \cdot \frac{21}{64} \cdot (1 - \frac{21}{64})} = 3,756 > 3$  erfüllt ist, kann die Binomialverteilung durch die Normalverteilung approximiert werden. Für die Stichprobe ergibt sich somit ein  $z$ -Wert von 1,86 welcher größer ist als der kritische Wert von 1,65. Die Abweichung des Gruppenmittels der Nachbefragung von dem der Vorbefragung ist somit signifikant. Folglich bewirkt das Training mit physiologischem Feedback im Gegensatz zu dem Training ohne Feedback eine deutliche Steigerung der Anwendung von Emotionsregulationsstrategien.

### 7.2.5 Transfer-Effekte auf andere Aufgaben

Nachfolgend soll untersucht werden, inwiefern sich das Training der Emotionsregulation durch das Spielen des Spiels mit physiologischem Feedback auf andere Aufgaben auswirken. Hierzu wird zunächst das Auction Game Diagnostic analysiert, welches vom Spielprinzip ähnlich dem Auction Game aufgebaut ist. Zusätzlich wird mittels des emotionalen Stroop-Tests das

Verhalten bei Aufgaben mit emotionalen Inhalten betrachtet. Zur Untersuchung der Transfer-Effekte wurde in allen Versuchspersonengruppen eine einheitliche Version der Tests verwendet.

Wie einführend erwähnt, wurden beide Tests jede Woche im Anschluss an die Emotionsregulationsintervention durchgeführt. Nachfolgende Untersuchungen beziehen sich jedoch ausschließlich auf die Ergebnisse aus der dritten Woche, da im Fokus der Untersuchungen die Transfer-Effekte im Allgemeinen, jedoch nicht zeitliche Lerneffekte in den Transfer-Aufgaben, stehen. Für die Analysen werden nur Daten von Versuchspersonen berücksichtigt, welche an allen drei Sitzungen teilgenommen haben und für welche während des Auction Games vollständige physiologische Daten aufgezeichnet wurden.

### 7.2.5.1 Auction Game Diagnostic

Im Gegensatz zum eigentlichen (didaktischen) Auction Game, welches auf eine Verbesserung der Emotionsregulationsfähigkeiten abzielt, dient die diagnostische Variante der Untersuchung von Verhaltens- und Erregungsmustern, welche sich während des Spiels zeigen.

Da die Versuchspersonen in der AUC-Bedingung bereits bei der didaktischen Variante des Auction Games eine leicht bessere Performance aufgewiesen haben, wird vermutet, dass sich dies auch im Auction Game Diagnostic zeigt, zumal die Spielschwierigkeit hier für beide Gruppen gleich ist. Des Weiteren wird angenommen, dass das physiologische Feedback während des Auction Game Didactic dazu führt, dass die mittlere Herzrate während des Spiels im Mittel für die Versuchspersonen der AUC-Bedingung niedriger ausfällt, als für die Kontrollgruppe, da sich auch dieser Effekt bereits in der didaktischen Variante gezeigt hat. Aufgrund fehlender Daten für zwei Versuchspersonen aus der AUC-Bedingung liegt die Stichprobengröße für die Analysen bei  $N_{CNT} = 23$  und  $N_{AUC} = 18$ .

Wie vermutet, zeigt sich für die Versuchspersonen der AUC-Bedingung eine leicht bessere Performance, als für die Kontrollgruppe (81,27 % vs. 81,01 %). Ein einseitiger  $t$ -Test zeigt jedoch, dass diese Unterschiede nicht signifikant sind ( $t(39) = 0,080$ ;  $p = 0,477$ ). Bei der relativen Veränderung der Herzrate im Vergleich zur Baseline sind die erwarteten Auswirkungen des Trainings deutlicher sichtbar: Während die mittlere Herzrate bei Versuchspersonen der AUC-Bedingung im Schnitt 1,18 % niedriger ist, als

während der Baseline, steigt diese für die Kontrollgruppe während des Spiels leicht an (2,56%). Ein einseitiger  $t$ -Test bestätigt die Annahme, dass die Herzratenveränderung unter der AUC-Bedingung signifikant geringer ausfällt, als bei der Kontrollgruppe ( $t(39) = -1,839$ );  $p < 0,05$ ). Es kann somit festgestellt werden, dass das Training zwar nicht zu Performanceunterschieden zwischen den beiden Versuchsgruppen führt, jedoch signifikante Unterschiede bei der Veränderung der Herzrate auftreten.

### 7.2.5.2 Emotionaler Stroop-Test

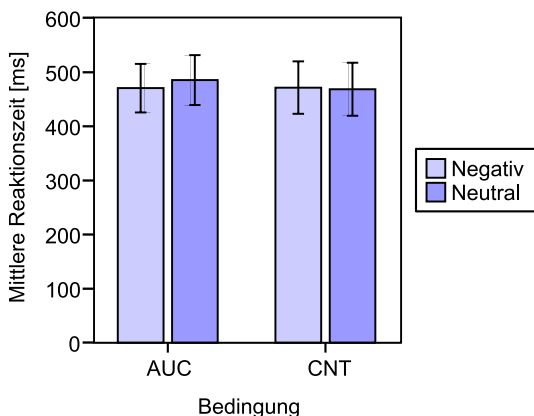
Ziel beim Auction Game Diagnostic war die Überprüfung, wie sich die gelernten Strategien zur Emotionsregulation in einem ähnlichen Spielszenario auswirken. Mittels des emotionalen Stroop-Tests soll nun untersucht werden, inwiefern es Unterschiede zwischen den beiden Gruppen bei der Bearbeitung von Aufgaben mit emotionalen Inhalten gibt.

Der emotionale Stroop-Test geht zurück auf Mathews & MacLeod (1985). Es handelt sich hierbei um eine Variation des klassischen Stroop-Tests, welcher bereits in Abschnitt 2.2.3.2 eingeführt wurde. Im Gegensatz zum klassischen Stroop-Test muss bei dem emotionalen Stroop-Test die Farbe von negativen, neutralen oder positiven Wörtern bestimmt werden. Es hat sich gezeigt, dass die Reaktionszeit bei negativen Wörtern höher ist, als bei anderen Wörtern. Dies kann darauf zurückgeführt werden, dass das negative Wort die Aufmerksamkeit von der eigentlichen Aufgabe der Farbbestimmung ablenkt (Gootjes et al., 2011).

In der vorliegenden Studie wurden den Versuchspersonen Wörter mit neutralen und negativen Inhalten präsentiert. Die Durchführung des emotionalen Stroop-Tests erfolgte jede Woche direkt im Anschluss an die didaktische Version des Auction Games. Aufgrund des Emotionsregulationstrainings, welches unter der AUC-Bedingung erfolgt ist, wird vermutet, dass die mittlere Differenz der Reaktionszeiten auf neutrale Wörter zu den Reaktionszeiten auf negative Wörter für Versuchspersonen in der AUC-Bedingung geringer ausfällt, als in der Kontrollgruppe.

Zur Überprüfung dieser Annahme werden die Daten aller Versuchspersonen aus der AUC- und CNT-Bedingung herangezogen, für welche in allen drei Wochen der Studie vollständige physiologische Informationen vorlagen. Es gilt daher  $N_{CNT} = 23$  und  $N_{AUC} = 20$ .

Es zeigt sich, dass die durchschnittlichen Reaktionszeiten für die Kontroll-



**Abbildung 7.7:** Reaktionszeiten für beide Gruppen für neutrale und negative Wörter (Fehlerbalken geben das 95 %-Konfidenzintervall an)

gruppe mit 470 ms kürzer ausfallen als die der Versuchsgruppe (478 ms). Wie bereits bei dem Auction Game Diagnostic, zeigt sich somit auch hier eine leicht niedrigere Gesamtperformance der Versuchsgruppe gegenüber der Kontrollgruppe. Interessant ist jedoch, dass die Reaktionszeit bei negativen Wörtern für die Versuchsgruppe niedriger ist als für neutrale Wörter, während bei der Kontrollgruppe die Reaktionszeiten für negative Wörter – wie es bei dem emotionalen Stroop-Test üblicherweise der Fall ist – leicht höher ausfallen, als für neutrale Wörter. Die durchschnittlichen Reaktionszeiten auf neutrale und negative Worte für beide Gruppen ist in Abbildung 7.7 veranschaulicht.

Eine einfaktorielle ANOVA mit dem Faktor Versuchsbedingung und den unabhängigen Variablen mittlere Reaktionszeit auf negative Wörter, mittlerer Reaktionszeit auf neutrale Wörter sowie der Differenz der Reaktionszeiten auf negative Wörter und neutrale Wörter zeigt, dass für die Versuchsgruppe und die Kontrollgruppe keine signifikanten Unterschiede für die mittleren Reaktionszeiten auf negative ( $F(1; 41) = 0,001$ ;  $p = 0,979$ ) und neutrale ( $F(1; 41) = 0,269$ ;  $p = 0,607$ ) Wörter bestehen. Jedoch weisen die beiden Gruppen einen signifikanten Unterschied ( $F(1; 41) = 5,694$ ;  $p < 0,05$ ) hinsichtlich der Differenz zwischen beiden Reaktionszeiten auf.

Der Unterschied zwischen den Reaktionszeiten für neutrale und negative Wörter fällt für die AUC-Gruppe somit signifikant geringer aus, als für die Kontrollgruppe. Dies kann als Indiz dafür gesehen werden, dass die während des Auction Game Didactic gelernten Emotionsregulationsstrategien von den Versuchspersonen der AUC-Gruppe auch in anderen Aufgaben angewendet werden. Eine mögliche Begründung für die insgesamt leicht langsameren Reaktionszeiten während des emotionalen Stroop-Tests kann darin gesehen werden, dass das vorherige Spielen des Auction Game Didactic für die AUC-Gruppe schwieriger war als für die Kontrollgruppe und die hiermit verbundene Anstrengung sich auch auf die nachfolgende Aufgabe ausgewirkt hat.

### 7.3 Zusammenfassung der Ergebnisse

Die Ergebnisse der vorherstehenden Analysen zeigen, dass sich beim Vergleich der Gruppen mit und ohne physiologischem Feedback deutliche Unterschiede sowohl bei der Leistung im Spiel als auch bei der Emotionsregulation ergeben. Insgesamt fiel die Steigerung der Spielperformance über den Verlauf der drei Wochen, in welcher das Spiel gespielt wurde, für die Kontrollgruppe etwas höher aus als für die AUC-Gruppe. Dies war jedoch nicht primäres Anliegen des Trainings. Vielmehr sollten mit Hilfe des Spiels die Emotionsregulationsfähigkeiten trainiert werden. Dies spiegelte sich in signifikanten Unterschieden der HRV während des Spiels sowie des Suppression-Wertes in Bezug auf das Spiel wider. Auch in der Nachbefragung drei Wochen nach dem Spiel zeigte sich, dass Versuchspersonen der ER-Gruppe häufiger Emotionsregulationsstrategien anwenden. Zwar ist der Unterschied zwischen beiden Gruppen nicht signifikant, dies kann jedoch auf die verhältnismäßig kurze Trainingsdauer zurückgeführt werden.

Auch sind nach der dritten Studienwoche Transfer-Effekte auf andere Aufgaben beobachtbar. Beim Spielen des Auction Game in der didaktischen Version zeigen sich zwar keine Performanceunterschiede, jedoch ist die Herzrate bei der Versuchsgruppe signifikant niedriger als bei der Kontrollgruppe. Zudem fällt bei dem emotionalen Stroop-Test die Differenz zwischen den Reaktionszeiten auf neutrale und negative Wörter bei der Versuchsgruppe geringer aus als bei der Kontrollgruppe, was auf die An-



wendung der zuvor erlernten Emotionsregulationsstrategien zurückgeführt werden kann.

Zusammenfassend lässt sich somit feststellen, dass das physiologische Feedback im Spiel einen klar erkennbaren Einfluss auf den Lernverlauf der Versuchspersonen hatte.

# 8 Integration von physiologischem Feedback in ein Arbeitsgedächtnistraining

In Kapitel 7 wurde eine Möglichkeit vorgestellt, wie der Nutzer lernen kann, sein emotionales Erregungsniveau dem Lernprozess anzupassen. Die Arbeiten in diesem Kapitel zielen nun darauf ab, eine Anwendung zu entwickeln, welche in der Lage ist, sich an das emotionale Erregungsniveau eines Nutzers anzupassen.

Bei der Entwicklung eines affektiven Lernsystems stehen nach D’Mello et al. (2008) neben der Fragestellung, welche emotionalen Zustände für einen Lernprozess relevant sind, vor allem die Aspekte im Vordergrund, wie diese Zustände möglichst gut erkannt werden, sowie wie hierauf adäquat reagiert werden kann. In Abschnitt 8.1 wird daher zunächst darauf eingegangen, wie genau diese Aspekte im Rahmen des Arbeitsgedächtnistrainings adressiert werden. Hierauf folgt eine Beschreibung der konkreten Ausgestaltung der Adaptionsmechanismen. Abschnitt 8.2 widmet sich anschließend der Analyse des entwickelten Systems.

Teile dieses Kapitels wurden bereits in Schaaff (2013) veröffentlicht.

## 8.1 Anwendungsdesign

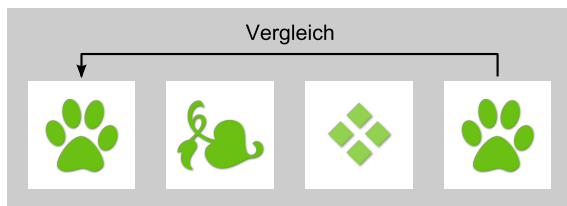
Wie in Abschnitt 2.2.3.2 dargestellt, gibt es eine Reihe standardisierter Verfahren zum Training des Arbeitsgedächtnisses. Trainingsprogramme, bei denen sich die Schwierigkeit im Laufe des Trainings verändert, führen hierbei zu größeren Trainingserfolgen, als Trainings mit konstantem Schwierigkeitsniveau (Klingberg, 2010). Die Anpassung der Trainingsschwierigkeit basiert hierbei in der Regel auf früheren Trainingsergebnissen – sei es aus einem Einstufungstest vor dem Training (z. B. Schmiedek et al., 2010)

oder während des Trainings (z. B. Heinzl et al., 2013). Der emotionale Zustand des Nutzers hingegen wird bei bisherigen Systemen bei der Anpassung der Trainingsschwierigkeit nicht betrachtet. Auch dass die Leistung eines Nutzers in Abhängigkeit des emotionalen Erregungsniveaus starken Schwankungen unterlegen ist (siehe Abschnitt 2.2.2) und somit zu verschiedenen Zeitpunkten unterschiedlich hohe Aufwände für das Erreichen eines bestimmten Ergebnisses erforderlich sind, bleibt hierbei unberücksichtigt.

Im Rahmen dieser Arbeit wird ein Arbeitsgedächtnistraining vorgestellt, bei welchem eine Adaption der Trainingsschwierigkeit basierend auf dem emotionalen Erregungsniveau eines Nutzers durchgeführt wird. Hierzu wird die  $n$ -Back-Task verwendet, da bei dieser die Trainingsschwierigkeit flexibel variiert werden kann. Die Implementierung der Anwendung erfolgt in Java. Zur Detektion des emotionalen Erregungszustandes eines Nutzers wird psychophysiologische Sensorik genutzt. Diese wird über die die xAffect-Software an das Arbeitsgedächtnistraining angebunden.

### 8.1.1 Trainingsprinzip

Im Rahmen des Trainings erfolgt bei der  $n$ -Back-Task die Präsentation verschiedener Symbole. Die Aufgabe des Nutzers besteht darin zu bestimmen, inwiefern das aktuell gezeigte Symbol mit dem Symbol, welches  $n$  Elemente zuvor präsentiert wurde, übereinstimmt. Abbildung 8.1 veranschaulicht diesen Ablauf für den Fall  $n = 3$ .



**Abbildung 8.1:** Prinzip des entwickelten Arbeitsgedächtnistrainings für den Fall  $n = 3$

Das Arbeitsgedächtnistraining umfasst üblicherweise mehrere Sessions, welche sich auf einen längeren Zeitraum (z.B. mehrere Tage) verteilen. Eine Session wiederum besteht aus mehreren Runden. Die Dauer einer Runde wurde in der vorliegenden Arbeit in Anlehnung an die Literatur

auf 60 Sekunden festgelegt. Zwischen den einzelnen Runden gibt es eine Pause, deren Länge der Nutzer selbst bestimmen kann.

Bei der  $n$ -Back-Task gibt es verschiedene Parameter, welche zur Variation der Schwierigkeit während des Trainings genutzt werden können:

- Veränderung von  $n$
- Veränderung der Lures
- Veränderung von Präsentationszeit und Inter-Stimulus-Intervall (ISI)

Für eine genauere Erläuterung dieser Parameter sei auf Abschnitt 3.4.2 verwiesen.

In Anlehnung an Heinzel et al. (2013) wurden die Adaptionstufen in dieser Arbeit derart gestaltet, dass zum einen das ISI in Schritten von 200 ms von 1000 bis 1800 ms variiert wird. Zum anderen erfolgt eine Variation des Parameters  $n$ . Für jeden Wert von  $n$  wird bei steigender Schwierigkeit zunächst das ISI schrittweise bis auf 1000 ms verringert. Ist diese Untergrenze erreicht, erfolgt im nächsten Schritt eine Erhöhung des Parameters  $n$  um 1. Die Verringerung der Schwierigkeitsstufen erfolgt analog hierzu: Das ISI wird zunächst bis zu einem Wert von 1800 ms erhöht und im nächsten Schritt der Parameter  $n$  um eins verringert.

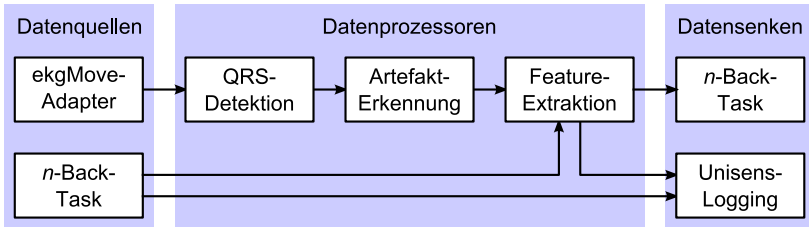
### 8.1.2 Integration des physiologischen Feedbacks

Bevor Informationen über das emotionale Erregungsniveau eines Nutzers in das Arbeitsgedächtnistraining integriert werden können, muss dieses zunächst detektiert werden. Neben der Eignung der ausgewählten Sensorik zur Detektion von emotionaler Erregung ist hierbei auch ein hoher Tragekomfort des Systems von Bedeutung, da der Nutzer durch die physiologische Sensorik möglichst wenig belastet werden soll.

In Kapitel 5 wurden verschiedene physiologische Signale hinsichtlich ihrer Eignung für die Erregungserkennung untersucht. Die Ergebnisse zeigten, dass basierend auf dem EKG-Signal eine gute Trennung unterschiedlicher Erregungszustände möglich ist. Zwar konnte das Ergebnis durch Kombination mehrere Signale leicht verbessert werden, jedoch erfordert dies auch das Tragen zusätzlicher Sensoren. Um die Belastung durch die physiologische Sensorik für den Nutzer möglichst gering zu halten, wird die Erkennung der emotionalen Erregung in dem dargestellten

System ausschließlich basierend auf dem EKG-Signal durchgeführt. Zur Aufzeichnung des EKG-Signals kommt erneut der ekgMove-Sensor zum Einsatz, da dieser hohen Tragekomfort mit guter Signalqualität sowie der Möglichkeit zur Online-Analyse der Daten verbindet.

Die Erregungserkennung wird mittels der xAffect-Software durchgeführt. Das hierfür verwendete Setup ist in Abbildung 8.2 dargestellt.



**Abbildung 8.2:** Setup der xAffect-Software für das adaptive Arbeitsgedächtnistraining

Der ekgMove-Sensor ist über den entsprechenden Adapter in die xAffect-Software eingebunden. Basierend auf dem EKG-Signal erfolgt eine kontinuierliche Detektion der R-Zacken. Nach einer Artefakterkennung werden hieraus die Parameter Herzrate und RMSSD extrahiert. Das Arbeitsgedächtnistraining fungiert in diesem Setup als Datenquelle und -senke zugleich. Als Datenquelle ist es für die Steuerung der Berechnung der physiologischen Parameter zuständig. Jeweils zu Beginn der Baseline-Phase bzw. der jeweiligen Runden wird ein Trigger an die Feature-Komponente gesendet. Diese startet daraufhin mit der Berechnung der Features. Das Ergebnis dieser Feature-Berechnung wiederum wird zum einen über den Unisens-Logger gespeichert, zum anderen der Trainingsanwendung im Anschluss an jede Runde zur Weiterverarbeitung zur Verfügung gestellt.

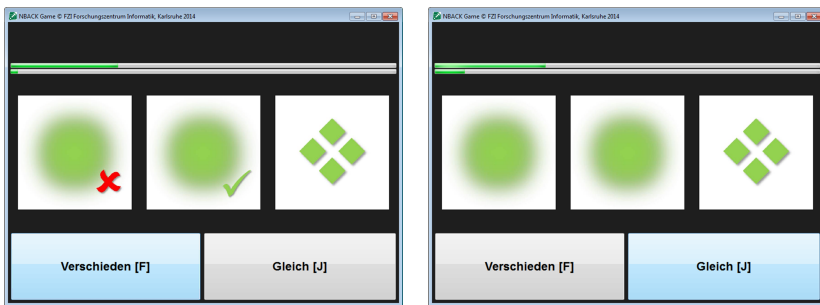
### 8.1.3 Konzeptionsstudie zur Feedbackanzeige

In den meisten Anwendungen, bei welchen die *n*-Back-Task zum Einsatz kommt, erhält der Nutzer üblicherweise direkt Feedback darüber, ob seine Zuordnung, dass ein Element *n* Elemente zuvor bereits aufgetaucht ist, richtig war. Während früher Anwendungstests fiel auf, dass das direkte Feedback auf einzelne Nutzer störend wirkte bzw. von der eigentlichen

Aufgabe ablenkte. Aus diesem Grund wurde im Rahmen einer Vorabanalyse untersucht, ob es Unterschiede im Lernverhalten bei der  $n$ -Back-Task in Abhängigkeit von der direkten Feedbackanzeige für die einzelnen Elemente gibt.

An der Studie nahmen insgesamt 46 Versuchspersonen teil. Ein Teilnehmer wurde nachträglich von den Analysen ausgeschlossen, da dieser angab, an einer Konzentrationsschwäche zu leiden. Von den verbleibenden Versuchspersonen entfallen 22 auf die Gruppe mit direktem Feedback, 23 auf die Gruppe ohne Feedback. 28 der Versuchspersonen waren männlich, 17 weiblich. Das Durchschnittsalter lag bei 26,36 Jahren ( $SD = 5,54$ ).

Für die Evaluation wurden zwei unterschiedliche Versionen der Software verwendet: eine Version mit und eine ohne direktem Feedback. Abbildung 8.3 zeigt die beiden unterschiedlichen Versionen. Bei der Version mit



**Abbildung 8.3:** Versionen der  $n$ -Back-Task zur Evaluation der Feedbackanzeige. Links: mit direktem Feedback, rechts: ohne direktes Feedback

direktem Feedback erhält der Nutzer direkt nach der Zuordnung, ob ein Element bereits  $n$  Elemente vorher angezeigt wurde, Feedback darüber, ob dies richtig (grüner Haken) oder falsch (rotes Kreuz) war. In der Version ohne direktes Feedback fehlen diese Elemente. Die Adaption erfolgt bei beiden Versionen in Abhängigkeit von der Spielperformance. Hierbei werden in Anlehnung an Heinzl et al. (2013) die Untergrenze von 60 % für eine Verringerung der Schwierigkeit und 80 % für eine Erhöhung der Schwierigkeit verwendet. Physiologische Daten sind für diese Evaluation nicht eingebunden.

Im Rahmen der Evaluation wurden die Teilnehmer gebeten, jeweils eine Session des Spiels mit 5 Runden zu spielen. Für die Analyse, inwiefern die Feedbackanzeige einen Unterschied auf die Spielerperformance hat, wurde die Veränderung der Schwierigkeitsstufen herangezogen, welche im Verlauf des Spiels durchlaufen wurden. Innerhalb der 5 Runden können sich die Spieler um maximal 4 Schwierigkeitsstufen verschlechtern oder verbessern.

In der Gruppe ohne direktes Feedback erfolgte eine durchschnittliche Verbesserung um 1,83 Level (SD = 1,61). In der Gruppe mit Feedback hingegen lag diese bei lediglich 0,64 Levels (SD = 1,94). Ein Wilcoxon-Rangsummentest zeigt, dass die Unterschiede zwischen den beiden Gruppen signifikant sind ( $W = 419,5$ ,  $Z = -1,991$ ,  $p < 0,05$ ). Somit fällt die Verbesserung in der Gruppe, die kein direktes Feedback angezeigt bekommt, signifikant größer aus, als bei der Gruppe, welche das Feedback angezeigt bekommt. Für alle nachfolgenden Untersuchungen wird daher eine Version ohne direktes Feedback verwendet.

### 8.1.4 Entwicklung des Adaptionalgorithmus

Ziel bei der Adaption des Systems an das emotionale Erregungsniveau des Nutzers ist es, diesen nach Möglichkeit während des gesamten Trainings auf einem optimalen Erregungsniveau zu halten. Hierdurch sollen sowohl Über- als auch Unterforderung vermieden werden.

Nachfolgender Abschnitt dient der Darstellung, wann und wie die Adaption im Rahmen des Arbeitsgedächtnistrainings erfolgt. Hierbei wird zunächst das allgemeine Adaptionkonzept vorgestellt. Zur Entwicklung des Adaptionalgorithmus wird im Anschluss daran eine Kalibrierungsstudie durchgeführt, um Informationen über den Verlauf der Herzaktivität während des Trainings zu erhalten. Basierend auf diesen Daten wird im nächsten Schritt ein System entwickelt, welches in der Lage ist, sowohl vor, während als auch nach dem Training den Nutzerzustand zu berücksichtigen.

#### 8.1.4.1 Adaptionkonzept

Grundsätzlich gibt es drei mögliche Zeitpunkte, zu welchen sich die Anwendung an den aktuellen Nutzerzustand anpassen kann: zu Beginn, während und nach einem Training. Alle drei Zeitpunkte sollen im Rahmen dieser Arbeit berücksichtigt werden und sind nachfolgend kurz beschrieben.

### Zu Beginn des Trainings

Wird bei herkömmlichen Arbeitsgedächtnistrainings eine Adaption zu Beginn vorgenommen, so erfolgt diese basierend auf der Performance in früheren Trainings. Der aktuelle Nutzerzustand bleibt hierbei unberücksichtigt. Dies kann schnell zu Frustration bei dem Nutzer führen, sofern das Training auf einem zu schwierigen Niveau beginnt, welches ihn überfordert, oder aber auch auf einem zu einfachen Niveau, das ihn langweilt. Im schlimmsten Fall kann dies sogar zu einem Abbruch des Trainings führen. Das hier vorgestellte Arbeitsgedächtnistraining nutzt Informationen aus einer dem Training vorausgehenden Baseline-Phase, um direkt zu Beginn der ersten Runde die Schwierigkeit an den aktuellen Nutzerzustand anzupassen.

### Während des Trainings

Arbeitsgedächtnistrainings setzen sich in der Regel aus mehreren kurzen Runden zusammen, welche nicht länger als 2 Minuten dauern. Auch das hier vorgestellte Training ist nach diesem Prinzip aufgebaut. Eine Adaption der Spielschwierigkeit innerhalb einer Runde kann – insbesondere wenn der Parameter  $n$  betroffen ist – leicht zu Verwirrung und / oder Ablenkung des Nutzers führen und sich somit negativ auf die Trainingsperformance auswirken. Aus diesem Grund wird in Anlehnung an die in der Literatur vorhandenen Ansätze auf eine Anpassung der Schwierigkeit während der einzelnen Runden verzichtet und die Adaption lediglich zwischen den Runden vorgenommen.

Auf eine Anzeige des emotionalen Erregungszustands während des Trainings wird bewusst verzichtet, da dies Emotionsregulationsprozesse während des Trainings anstoßen könnte. Da Emotionsregulation jedoch in den gleichen Gehirnarealen vonstatten geht, die auch vom Arbeitsgedächtnis verwendet werden, könnten bewusste Emotionsregulationsprozesse zu einer Beeinträchtigung der Arbeitsgedächtnisleistung führen.

### Am Ende des Trainings

In der Regel wird der Trainingsfortschritt im Vergleich zu den vorhergehenden Runden bzw. Sessions präsentiert. Jedoch kann der Aufwand für die gleiche Leistung im Training in Abhängigkeit des Erregungszustands unterschiedlich groß sein. Wird der Erregungszustand während



eines Trainings in die Präsentation des Lernfortschritts integriert, kann diesen Unterschieden Rechnung getragen werden.

#### 8.1.4.2 Kalibrierungsstudie

Vor dem Hintergrund, einen möglichst validen Adaptionsalgorithmus basierend auf den physiologischen Daten zu entwickeln, wurde im Rahmen dieser Arbeit eine Kalibrierungsstudie durchgeführt. An dieser Studie nahmen insgesamt 26 Versuchspersonen teil. Von diesen waren 22 männlich und 4 weiblich. Das Durchschnittsalter der Versuchspersonen lag bei 27,15 Jahren ( $SD = 3,77$ ).

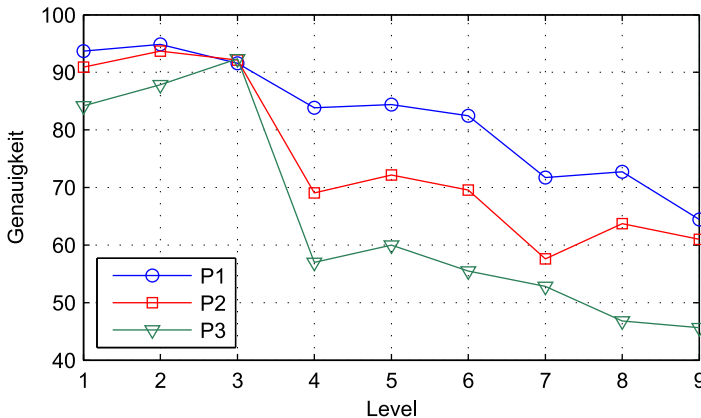
Im Rahmen der Kalibrierungsstudie wurde eine Version des Trainings verwendet, bei welcher die Trainingsschwierigkeit von Runde zu Runde schwieriger wird. Die Konfiguration des ISI sowie des Parameters N für die jeweiligen Runden sind in Tabelle 8.1 dargestellt.

Runde	ISI	$n$
1	1800	1
2	1400	1
3	1000	1
4	1800	2
5	1400	2
6	1000	2
7	1800	3
8	1400	3
9	1000	3

**Tabelle 8.1:** Konfiguration der Trainingsschwierigkeit im Rahmen der Kalibrierungsstudie

Zur besseren Analyse des Zusammenhangs zwischen der Herzratenveränderung und den Erregungsniveaus werden die Versuchspersonen in drei Gruppen eingeteilt, die auf ihren Quartilsrängen hinsichtlich der Durchschnittsperformance während des Kalibrierungstrainings basieren. Die obere Gruppe (P1) bilden alle Versuchspersonen, deren durchschnittliche Performance über dem  $Q_{0,75}$ -Wert liegt. Die mittlere Gruppe (P2) besteht aus allen Person, deren Performance zwischen dem  $Q_{0,75}$ - und dem  $Q_{0,25}$ -Quartil liegt und die untere Gruppe (P3) aus allen Personen mit einer Performance unter dem  $Q_{0,25}$ -Wert. In Abbildung 8.4 ist die durchschnitt-

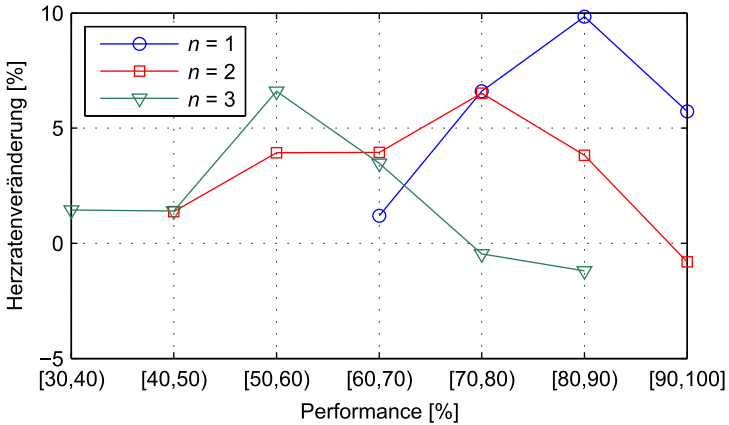
liche Performance der drei Gruppen im Verlauf über die einzelnen Runden der Kalibrierungsphase dargestellt. Während bei der Schwierigkeitsstufe  $n = 1$  in Runde 1 bis 3 die Leistungen aller drei Gruppen noch verhältnismäßig dicht beieinander liegen, findet ab Runde 4 bei Gruppe P3 ein deutlicher Leistungseinbruch auf unter 60 % statt. Im weiteren Verlauf geht die Leistung weiterhin leicht zurück, ähnlich starke Leistungseinbrüche sind jedoch nicht zu beobachten. Bei der Gruppe P2 sind beide Erhöhungen des Parameters  $n$  deutlich in der Performance zu erkennen. Bei  $n = 2$  sinkt die Performance zunächst auf Werte um rund 70 % und geht bei  $n = 3$  auf rund 60 % zurück. In Gruppe P1 hingegen zeigt sich der Übergang von  $n$  auf 2 nur durch einen leichten Leistungsrückgang. Erst bei  $n = 3$  geht die Performance auf Werte um 70 % zurück.



**Abbildung 8.4:** Mittlere Trainingsperformance pro Runde nach Quartilsrängen

Interessant sind die Unterschiede hinsichtlich der Korrelation zwischen der Performance und der Herzratenveränderung zwischen den einzelnen Gruppen. In Gruppe P3 zeigt sich ein signifikanter positiver Zusammenhang zwischen der mittleren Herzratenveränderung im Vergleich zur Baseline ( $r(54) = 0,354$ ;  $p < 0,01$ ). Auch in Gruppe P2 ist ein positiver Zusammenhang zu beobachten ( $r(117) = 0,179$ ;  $p = 0,054$ ). Für P1 hingegen zeichnet sich ein leicht negativer Zusammenhang ab ( $r(63) = -0,077$ ;  $p = 0,550$ ). Bei P3 ist eine höhere Performance somit mit einer höheren

Herzratensteigerung verknüpft wohingegen der Zusammenhang bei P2 geringer ausfällt bzw. bei P1 sogar umgekehrt ist. Der positive Zusammenhang in P3 lässt sich damit erklären, dass die Versuchspersonen bereits zu einem sehr frühen Spielzeitpunkt von der Spielschwierigkeit überfordert waren. Laut Befragung der Personen dieser Gruppe stellten sich durch die Überforderung schnell Frustration und Resignation ein, welche mit einem niedrigen emotionalen Erregungsniveau verbunden ist. In Abbil-



**Abbildung 8.5:** Relative Herzratenveränderung in Abhängigkeit von der Trainingsperformance

dung 8.5 ist die mittlere relative Herzratenveränderung zur Baseline im Vergleich zur Trainingsperformance dargestellt. Hierbei wird differenziert nach dem Wert des Parameters  $n$ , da dieser die Trainingsschwierigkeit wesentlich beeinflusst. Es zeigt sich für alle Werte von  $n$  ein parabelartiger Kurvenverlauf. Wie erwartet, verschiebt sich die Kurve mit zunehmender Trainingsschwierigkeit nach links: Bei höherer Schwierigkeit tritt schon bei geringerer Performance eine hohe Herzrate auf. Hohe Performancewerte, werden nur von Personen erzielt, denen das Training insgesamt leichter fällt. Hierdurch wird weniger emotionale Erregung erzeugt, als wenn das Training sie voll fordern würde, was zu dem Absinken der durchschnittlichen Herzratenveränderung mit zunehmender Leistung führt. Auch für das Absinken der Performance ist ein Rückgang der durchschnittlichen Herzratenveränderung zu beobachten. Die Ursache hierfür kann darin gesehen

werden, dass viele der Versuchspersonen ab einem gewissen Schwierigkeitsniveau resigniert aufgeben bzw. Frustration zeigen. Frustration und Resignation führen jedoch zu einer niedrigeren emotionalen Erregung im Vergleich zu einer hohen Auslastung durch das Training.

Für den Adaptionalgorithmus bedeutet dies, dass eine niedrige Herzrate nicht nur darauf hindeuten kann, dass eine Person unterfordert ist, sondern auch auf das Gegenteil: dass eine Person aufgrund von Überforderung bzw. Frustration aufgegeben hat, sich weiter anzustrengen. Aus diesem Grund muss bei dem hier vorgestellten Training neben der Herzrate auch die Trainingsperformance für den Adaptionprozess mit berücksichtigt werden.

Bei dem vorgestellten Algorithmus werden die Forschungsergebnisse von Baker et al. (2010) mit einbezogen, nach welchen Frustration beim Lernen als weniger kritisch einzustufen ist als Langeweile. Da Frustration weniger beständig ist als Langeweile, ist es folglich weniger kritisch, wenn die Versuchspersonen kurzfristig überfordert sind, als wenn es während des Trainings zu Unterforderung kommt.

#### 8.1.4.3 Adaption zu Trainingsbeginn

Für die Adaption zu Beginn der jeweiligen Trainingssession wurden die Daten aus der in Kapitel 5 vorgestellten Studie herangezogen. Basierend auf diesen Daten wurde mittels der LDA eine Funktion ermittelt, welche zur Trennung zwischen hoher und niedriger Erregung basierend auf den absoluten Werten des EKG-Signals genutzt werden kann. Als Fenstergröße wurde die Länge der vor dem Training verwendeten Baseline von 60 Sekunden verwendet. Aus dieser Analyse ergibt sich folgende Trennfunktion:

$$y = -6,2540 + 0,07 \cdot \text{HR} + 0,0031 \cdot \text{RMSSD} \quad (8.1)$$

wobei niedrigere Werte auf ein geringeres emotionales Erregungsniveau, höhere Werte auf ein höheres emotionales Erregungsniveau schließen lassen.

Basierend auf den Daten aus der Kalibrierungsstudie werden folgende Schwellwerte für die initiale Adaption festgelegt:

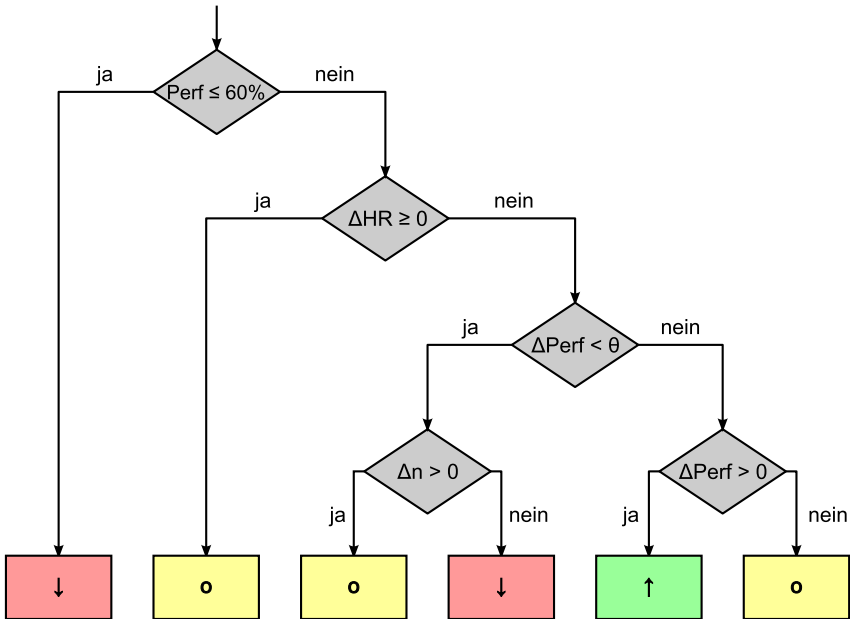
- $y < -0,79$ : Erhöhung der Trainingsschwierigkeit
- $y > 0,24$ : Verringerung der Trainingsschwierigkeit

Diese Schwellwerte wurden so gewählt, dass die Wahrscheinlichkeit für eine

Erhöhung der Trainingsschwierigkeit höher ist, als für eine Verringerung der Trainingsschwierigkeit, damit der Nutzer direkt von Beginn an möglichst stark gefordert ist. In Abbildung 8.6 ist der entwickelte Algorithmus für die Adaption der Trainingsschwierigkeit zwischen den einzelnen Runden dargestellt.

8.1.4.4 Adaption während des Trainings

Wie bereits erwähnt, erfolgt die Adaption während des Trainings basierend auf einer Kombination aus der Herzratenveränderung und der Trainingsperformance. Der Algorithmus ist so aufgebaut, dass dieser in erster Linie



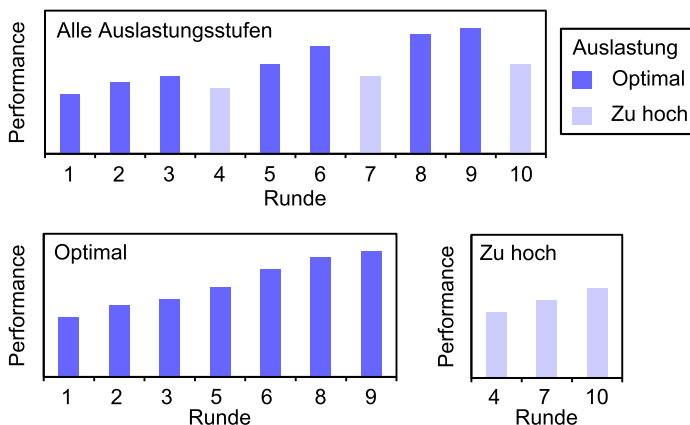
**Abbildung 8.6:** Algorithmus zur Adaption der Trainingsschwierigkeit (rot: Verringerung der Schwierigkeit, gelb: keine Veränderung der Schwierigkeit, grün: Steigerung der Schwierigkeit)

auf die Vermeidung von Über- und Unterforderung abzielt, welche sich wie oben dargestellt jeweils in einer fallenden Herzrate widerspiegeln.

Zur Differenzierung, ob Über- oder Unterforderung vorliegt, wird neben der Herzratenveränderung zur Vorrunde die Veränderung der Trainingsperformance im Vergleich zur Vorrunde herangezogen. Bei gesteigerter Performance kann davon ausgegangen werden, dass das Absinken der Herzrate auf Unterforderung zurückzuführen ist. Bei fallender Performance hingegen wird angenommen, dass das Absinken der Herzrate durch Überforderung zustande kommt. Zusätzlich wurde bei fallender Performance der Schwellwert  $\theta$  eingeführt, welcher sich in Abhängigkeit des jeweiligen ISIs ergibt:  $\theta$  entspricht hier der prozentualen Veränderung, die bei konstantem ISI durch Falschbeantworten zweier Elemente zustande kommen würde. Da eine Erhöhung des Parameters  $n$  kurzfristig zu größeren Schwankungen der Trainingsperformance führen kann, wurde zusätzlich eine Überprüfung eingebaut, ob  $n$  im Vergleich zur letzten Runde erhöht wurde. Ist dies der Fall, wird eine Verringerung der Trainingsschwierigkeit in der nächsten Runde verhindert um Oszillationen bei  $n$ -Übergängen auszuschließen. Zur Vermeidung zu großer Frustration wurde zudem ein Mechanismus implementiert, der bei Unterschreiten einer Performance von 60 % in einer Runde in der nächsten Runde die Trainingsschwierigkeit verringert.

#### 8.1.4.5 Adaption nach dem Training

Auch nach dem Training kann das emotionale Erregungsniveau des Nutzers bei der Präsentation der Trainingsergebnisse Berücksichtigung finden. Das Grundprinzip der hier verwendeten Ergebnispräsentation ist einfach: Sieht der Nutzer, dass er sich im Vergleich zum vorherigen Training verbessert hat, so kann sich dies auf zukünftige Trainings motivationssteigernd auswirken. Abbildung 8.7 zeigt die Unterschiede zwischen der normalen Ergebnispräsentation und der Präsentation in Abhängigkeit des Erregungsniveaus, welches sich in der Auslastung des Nutzers widerspiegelt. Während der Nutzer sich in dem Beispiel von Runde 9 auf Runde 10 etwas verschlechtert hat, ist in der auslastungsabhängigen Präsentation der Ergebnisse eine Verbesserung gegenüber der letzten Runde, bei welcher der Nutzer zu stark ausgelastet war zu erkennen.



**Abbildung 8.7:** Methoden zur Darstellung der Trainingsergebnisse. Oben: ohne Berücksichtigung der Auslastung, unten: mit Berücksichtigung der Auslastung

## 8.2 Systemanalyse

Bei der Evaluation des entwickelten Systems steht die Analyse des Adaptionmechanismus im Vordergrund. Zudem wird untersucht, inwiefern dieser bei den Versuchspersonen zu Unterschieden hinsichtlich des Trainingsverhaltens im Vergleich zu einer nicht-adaptiven Version führt. Neben der Verbesserung in der Aufgabe selbst werden hierbei auch motivationale Faktoren mit in die Untersuchungen einbezogen.

An der Studie nahmen insgesamt 20 Versuchspersonen teil, welche bereits zuvor an der Kalibrierungsstudie teilgenommen hatten. Von diesen waren 17 männlich und 3 weiblich mit einem Durchschnittsalter von 28,15 Jahren ( $SD = 3,50$ ). Die Versuchspersonen wurden zufällig auf zwei Gruppen aufgeteilt: Bei der ersten Gruppe kam die Version des Trainings zum Einsatz, bei welchem eine Adaption der Trainingsschwierigkeit mittels physiologischem Feedback erfolgt. Bei der Kontrollgruppe hingegen erfolgte eine initiale Adaption der Trainingsschwierigkeit basierend auf den Ergebnissen der Kalibrierungsstudie. Im Anschluss daran fand keine weitere Schwierigkeitsanpassung mehr statt.



**Abbildung 8.8:** Ablauf der Systemevaluation

Die Studie setzt sich aus mehreren Sessions zusammen. Der genaue Ablauf ist in Abbildung 8.8 dargestellt. Zu Beginn erfolgt eine Kalibrierung des Systems mittels der in Abschnitt 8.1.4 vorgestellten Version der  $n$ -Back-Task. Anschließend wird an 5 aufeinanderfolgenden Tagen je eine Trainingssession mit 10 Runden durchgeführt. Zum Abschluss wird mittels der Kalibrierungsversion der Trainingsfortschritt ermittelt. Im Anschluss an das Training wird den Versuchspersonen eine modifizierte Version des Game Experience Questionnaire (GEQ) vorgelegt. Hierbei wurde die In-Game-Version des Fragebogens ohne den Item des Vorstellungsvermögens verwendet. Zusätzlich erfolgt im Anschluss an die dritte Session eine Analyse der System-Usability anhand der System Usability Scale (SUS).

Aufgrund des relativ geringen Stichprobenumfanges erfolgen nachfolgende Analysen nur deskriptiver Art, da mittels inferenzstatistischer Verfahren keine aussagekräftigen Schlüsse möglich sind.

### 8.2.1 Adaptionsverhalten der Anwendung

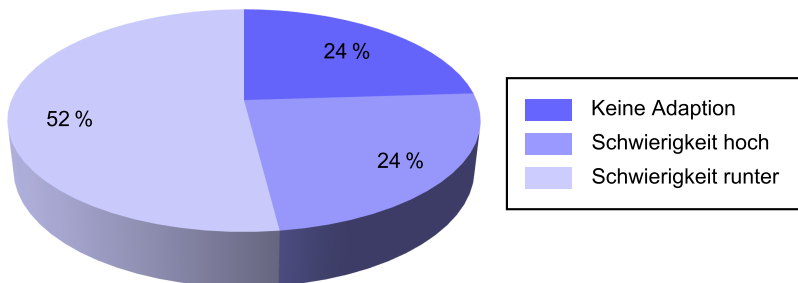
Ziel der Adaption ist es, den Nutzer zu jedem Zeitpunkt möglichst optimal zu fordern. Bei dem Entwurf des Adaptionsalgorithmus stand zudem die Erkenntnis im Vordergrund, dass Langeweile sich langfristig negativer auf den Lernerfolg ausübt als Frustration (Baker et al., 2010). Aus diesem Grund sollte die Adaption derart erfolgen, dass Langeweile durch Unterforderung nach Möglichkeit so selten wie möglich auftritt. Frustration durch Überforderung hingegen wurde als weniger kritisch eingestuft.



Nachfolgend soll nun untersucht werden, inwiefern die Adaption den Anforderungen genügt.

### Adaption vor Trainingsbeginn

Im ersten Schritt wird die Adaption zu Beginn der jeweiligen Trainingssessions untersucht. Diese wurde basierend auf den physiologischen Parametern während der Baseline vorgenommen. In Abbildung 8.9 ist dargestellt, wie häufig aufgrund der Vorgabe eine Adaption vorgenommen wurde. Es zeigt sich, dass die Trainingsschwierigkeit in den meisten Fällen zu Beginn des Trainings erhöht wurde. Von den 12 Fällen bei welchen zu Beginn des



**Abbildung 8.9:** Adaption zu Beginn des Spiels

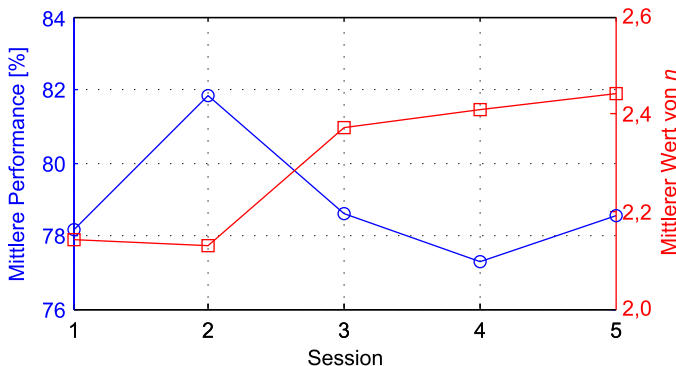
Training verringert wurde, wurde im Anschluss an die erste Runde in 11 Fällen keine weitere Adaption vorgenommen, in einem Fall wurde die Schwierigkeit in der nächsten Runde erhöht. Die zeigt, dass in fast allen Fällen durch die Verringerung der Trainingsschwierigkeit eine optimale Auslastung des Nutzers erreicht wurde. Bei den 12 Runden, in denen bei Trainingsbeginn keine Adaption vorgenommen wurde, wurde für die erste Runde in 8 Fällen eine zu geringe Nutzerauslastung festgestellt. Nur in 4 Fällen war die Auslastung optimal. Eine Erhöhung der Schwierigkeit hätte hier folglich zu einer besseren Auslastung führen können. Insgesamt wurde in 26 Runden eine initiale Erhöhung des Schwierigkeitsniveaus basierend auf den Vitalparametern vorgenommen. In 11 dieser Runden war die anschließende Nutzerauslastung optimal, in 13 Fällen zu niedrig und in 2 Fällen zu hoch. Die Erhöhung der Schwierigkeit war folglich in fast allen Fällen angemessen, hätte jedoch in Einzelfällen auch mehr als eine Schwierigkeitsstufe umfassen können.

### Adaption während des Trainings

Mittels der Adaption während des Trainings soll erreicht werden, dass der Nutzer möglichst selten den Zustand der Über- oder Unterforderung erreicht. Eine optimale Forderung spiegelt sich darin wider, dass die Genauigkeit während des gesamten Trainings ungefähr gleich bleibt. Dies kann durch eine Anpassung des Schwierigkeitsniveaus erzielt werden.

Insgesamt kommt es während während des Trainings in 20,89% aller Fälle zu einer Erhöhung der Schwierigkeit (Detektion von Unterforderung). In 61,33% der Fälle bleibt die Schwierigkeit gleich und in 17,78% der Fälle wird die Schwierigkeit gesenkt (Detektion von Überforderung).

In Abbildung 8.10 ist die Entwicklung der Performance sowie die Trainingsschwierigkeit über die einzelnen Sessions dargestellt. Es zeigt sich, dass bei der Performance zwar Schwankungen zwischen den einzelnen Sessions auftreten, insgesamt die durchschnittliche Performance der letzten Session jedoch nicht weitaus höher ist als in der ersten Session. Die Trainingsschwierigkeit hingegen – gemessen an dem durchschnittlichen Wert des Parameters  $n$  – steigt im Verlauf der Sessions immer weiter an. Somit trägt die kontinuierliche Adaption während des Trainings dazu bei, dass die Auslastung der Versuchspersonen während des Trainings konstant bleibt.



**Abbildung 8.10:** Verlauf von Performance und Schwierigkeitsgrad über die Sessions

Insgesamt fand im Laufe des Trainings eine Steigerung der Performan-

ce um lediglich 0,91 % statt. Diese verhältnismäßig geringe Steigerung ist durch die ständige Erhöhung des Schwierigkeitsgrades zu begründen: Von der ersten bis zur letzten Runde erhöhte sich die durchschnittliche Schwierigkeit um insgesamt 14,28 %. Obwohl also anhand der Performance keine Leistungssteigerung zu erkennen ist, drückt sich diese deutlich in der erhöhten Schwierigkeit aus.

### Adaption nach dem Training

Im Anschluss an das Training kann sich der Nutzer das Feedback über seine Leistung in Abhängigkeit seines Nutzerzustands anzeigen lassen. Hintergrund ist, dass der absolute Vergleich der Trainingsgenauigkeit zur Vorrunde zwar eine Verringerung darstellen kann, im Vergleich zu anderen Runden, in welchen der Nutzer in einem ähnlichen Zustand war, jedoch durchaus eine Steigerung bedeuten kann. Insbesondere in Phasen, in welchen durch das System eine Überforderung detektiert wurde, kann eine absolute Betrachtung demotivierend wirken.

Betrachtet man die durchschnittliche Veränderung der Genauigkeit der Runden, in welchen Überforderung detektiert wurde im Vergleich zur jeweils vorhergehenden Runde, so ergibt sich eine Veränderung von -13,89 %. Setzt man die Genauigkeit jedoch ins Verhältnis zur vorhergehenden Runde, in welcher bei einem Nutzer ebenfalls Überforderung detektiert wurde, so zeigt sich eine durchschnittliche Steigerung von +1,56 %. Umgekehrt verhält es sich für Runden, in welchen Unterforderung erkannt wurde: Im Vergleich zur Vorrunde zeigt sich hier eine durchschnittliche Steigerung von +11,69 %, im Vergleich zur vorhergehenden Runde, bei welcher Unterforderung vorlag, jedoch nur +0,48 %. Bei der optimalen Auslastung gibt es nur geringe Unterschiede zwischen dem Vergleich zur vorhergehenden Runde (+1,04 %) und dem Vergleich zu vorhergehenden Runden mit gleicher Auslastung (0,52 %).

Die Ergebnisse zeigen, dass sich insbesondere in Phasen von Überforderung eine nutzerzustandsabhängige Anzeige des Feedbacks positiv auf die Motivation eines Nutzers auswirken kann, da die Ergebnisse auf diese Weise positiver ausfallen. In Phasen von Unterforderung hingegen, wird der Trainingsfortschritt durch eine derartige Anzeige eher abgeschwächt, so dass hier eine nutzerzustandsabhängige Anzeige nicht zwangsläufig vorteilhaft ist.

### 8.2.2 Analyse des Trainingserlebnisses

Zur Analyse des Trainingserlebnisses wurde der GEQ-Fragebogen herangezogen, da dieser sehr gut die für diese Untersuchung relevanten Dimensionen dargestellt. In Abbildung 8.11 sind die Ergebnisse des GEQ für die beiden Versuchspersonengruppen dargestellt. Es wird ersichtlich, dass sich beide Versuchspersonengruppen während des Trainings im gleichen Maße kompetent fühlen. Dies ist überraschend, da die durchschnittliche Trainingsschwierigkeit in der Versuchsgruppe über der der Kontrollgruppe liegt. Jedoch kann dies auch als Maß dafür gesehen werden, dass die Versuchsgruppe durch die Veränderung der Schwierigkeit nicht überfordert war. Hinsichtlich des Flow-Erlebens fällt die Bewertung für die Versuchsgruppe wie erwartet etwas besser aus als für die Kontrollgruppe. Somit führt die Adaptivität des Trainings tatsächlich zu einer Steigerung des Flow-Erlebens. Dies zeigt sich auch bei der Anspannung, welche in der Versuchsgruppe niedriger ausfällt als für die Kontrollgruppe. Zudem wird das Training aufgrund der Adaption von der Versuchsgruppe insgesamt als herausfordernder empfunden als von der Kontrollgruppe. Insgesamt wird das Training von der Kontrollgruppe etwas negativer wahrgenommen als von der Versuchsgruppe, was sich in den Bewertungen des positiven

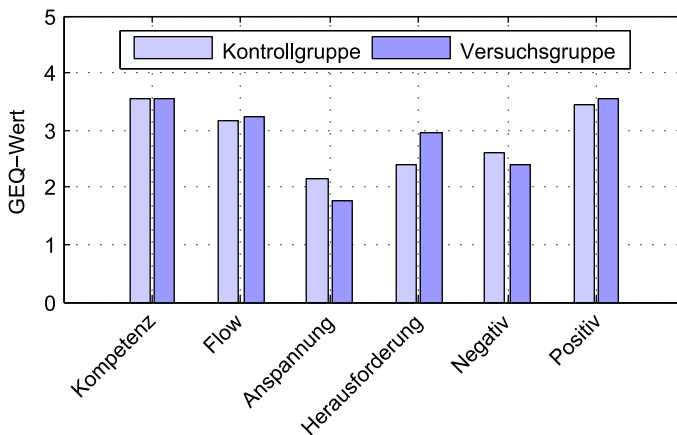


Abbildung 8.11: Ergebnisse des GEQ-Fragebogens

und negativen Affekts widerspiegelt.

Anhand des Fragebogens ist somit zu erkennen, dass die Kontrollgruppe insgesamt besser gefordert ist als die Versuchsgruppe, was ein positiveres Trainingserlebnis sowie ein besseres Flow-Empfinden zur Folge hat.

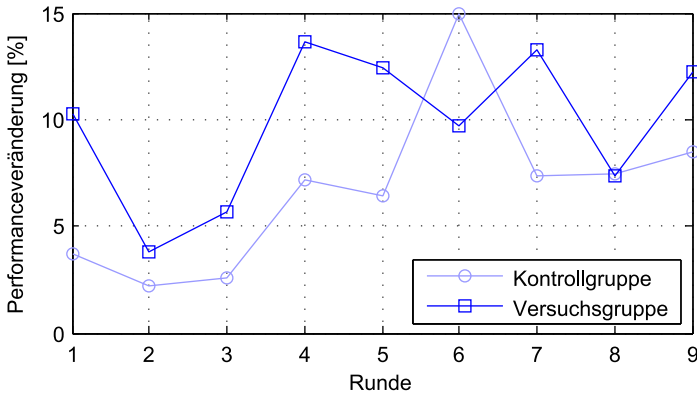
### 8.2.3 Analyse der Performanceveränderung

Zum Vergleich der Performanceveränderung für die beiden Gruppen durch das Training wurde die Verbesserung hinsichtlich der Performance von der Kalibrierungssession zur Abschlussession untersucht, da diese Sessions hinsichtlich der Schwierigkeit für beide Versuchspersonengruppen gleich sind.

Wie in Abbildung 8.12 dargestellt, zeigt sich für die Versuchsgruppe über nahezu alle Runden und somit auch über nahezu alle Schwierigkeitsstufen eine deutlich stärkere Leistungssteigerung als für die Kontrollgruppe. Während die durchschnittliche Performanceverbesserung für die Versuchsgruppe bei 9,82 % liegt, wird in der Kontrollgruppe lediglich eine Verbesserung von 6,72 % erzielt. Zudem ist zu beobachten, dass sich in der Herzrate sowie RMSSD für die Versuchsgruppe eine niedrigere Erregung widerspiegelt: Im Vergleich zur Kalibrierungssession ist in der Abschlussession die relative Herzratenveränderung zur Baseline 3,23 % niedriger, während für die Kontrollgruppe ein leichter Anstieg um 0,50 % zu beobachten ist. Für RMSSD hingegen nimmt die relative Veränderung zur Baseline für die Versuchsgruppe um 5,61 % zu während diese für die Kontrollgruppe um 11,30 % abnimmt. Anhand der Ergebnisse zeigt sich, dass die Versuchsgruppe der Kontrollgruppe hinsichtlich der Steigerung der Trainingsperformance deutlich überlegen ist.

### 8.2.4 Untersuchung der Systemusability

Ein wichtiger Aspekt bei dem Entwurf einer Lernanwendung ist eine hohe Benutzerfreundlichkeit der Anwendung. Aus diesem Grund erfolgte zusätzlich zu der eigentlichen Evaluation der Anwendung eine Analyse der Systemusability. Hierfür wurde die von Brooke (1996) entwickelte System Usability Scale (SUS) verwendet. Obwohl der Fragebogen im Vergleich zu anderen Usability-Fragebögen relativ einfach gehalten ist, kann mit diesem eine sehr hohe Reliabilität erzielt werden (Tullis & Stetson, 2004).



**Abbildung 8.12:** Performanceveränderung für beide Gruppen von Kalibrierungssession zu Abschlussession

Auf der Skala des Fragebogens können Werte von 0 bis 100 erzielt werden, wobei 100 der höchsten Usability entspricht. Im Durchschnitt wird ein Wert von etwa 70 vergeben, so dass ein System mit einem Wert über 70 als überdurchschnittlich eingestuft werden kann, während weniger als 50 Punkte als inakzeptabel gelten (Bangor et al., 2008).

Für die entwickelte Anwendung ergibt sich ein mittlerer Usability-Score von 78,63%. Mit diesem Wert liegt die Anwendung nach Bangor et al. (2008) im Vergleich zu anderen Anwendungen im oberen Quartil hinsichtlich der Usability und ist somit als gut bis sehr gut einzustufen.

## 8.3 Zusammenfassung der Ergebnisse

Ziel der vorstehenden Untersuchungen war die Integration von physiologischem Feedback in ein Arbeitsgedächtnistraining. In diesem Kapitel wurde auf Grundlage der Daten einer Kalibrierungsstudie ein Algorithmus entwickelt, welcher basierend auf den physiologischen Daten eines Nutzers die Schwierigkeit einer  $n$ -Back-Task an den Nutzerzustand anpasst. Der große Vorteil dieses Algorithmus ist darin zu sehen, dass die Adaption personenunabhängig erfolgt. Zudem basiert die Adaption während des Trai-

nings auf absoluten Werten, so dass eine vorhergehende Kalibrierungsphase nicht zwangsläufig erforderlich ist.

Im Rahmen einer Studie erfolgte ein Vergleich des entwickelten physiologischen Adaptionsmechanismus mit einer Version des Trainings, bei welcher die Adaption nur zu Beginn des Trainings basierend auf einem Einstufungstest erfolgt. Die Ergebnisse der durchgeführten Analysen zeigen, dass die entwickelten Adaptionsmechanismen zu einer sinnvollen Anpassung der Schwierigkeit während des Arbeitsgedächtnistrainings geführt haben. Es zeigte sich zudem, dass die Adaption basierend auf den physiologischen Parametern zu einer größeren Performanceverbesserung bei einer Kontrollaufgabe führte, als eine Adaption basierend auf einem vorhergehenden Einstufungstest. Zudem verdeutlichen die Ergebnisse des GEQ, dass die Wahrnehmung des Trainings bei der Versuchsgruppe weitaus positiver ausfällt als in der Kontrollgruppe.

# 9 Zusammenfassung und Ausblick

Emotionale Erregung übt einen entscheidenden Einfluss auf den Lernerfolg aus. Dies spiegelt sich in Modellen wie dem Yerkes-Dodson-Gesetz oder dem Flow-Modell wider. Während in der Mensch-Mensch-Kommunikation Emotionen wie selbstverständlich berücksichtigt werden, findet die Information über das emotionale Erregungsniveau in technologiegestützten Lernumgebungen derzeit nur unzureichend Beachtung. Diese Arbeit stellt einen Ansatz vor, wie emotionale Erregung mittels physiologischer Sensorik erfasst und sinnvoll in den Lernprozess integriert werden kann.

## 9.1 Zusammenfassung

Im Fokus dieser Arbeit stand die Integration von Informationen über den emotionalen Erregungszustand einer Person in Lernanwendungen. Damit eine Nutzung unter Alltagsbedingungen erfolgen kann, sollte diese Integration mit möglichst wenig zusätzlicher Belastung für den Nutzer verbunden sein. Aus diesem Grund wurde eine kraftsensitive Maus entwickelt, welche über den Tastendruck Informationen über den Erregungszustand einer Person reflektieren kann. Auch weitere physiologische Sensoren wurden in die Analysen mit einbezogen. Zur flexiblen Integration der Sensorsignale in die Lernanwendungen wurde zudem die xAffect-Software entwickelt – eine flexible Middleware, welche es ermöglicht, anderen Anwendungen basierend auf physiologischen Signalen Online-Feedback über emotionale Nutzerzustände zur Verfügung zu stellen.

Um eine möglichst valide Integration von Informationen über das emotionale Erregungsniveau einer Person in Lernanwendungen zu ermöglichen, wurde im Rahmen dieser Arbeit zunächst basierend auf psychophysiologischen Signalen untersucht, welche Signale am besten zur Erkennung emotionaler Erregungszustände geeignet sind. Hierzu wurde eine Studie durchgeführt, bei welcher den Versuchspersonen mittels eines Spiels hohe



bzw. niedrige emotionale Erregung induziert wurde. Es zeigte sich, dass bereits basierend auf dem EKG-Signal eine gute Erkennung der emotionalen Erregung erreicht werden kann. Zwar konnte durch die Hinzunahme weiterer Kanäle noch eine leichte Verbesserung der Erkennungsleistung erzielt werden, jedoch fiel diese verhältnismäßig gering aus. In einer weiteren zur Validierung der Ergebnisse durchgeführten Studie konnten diese Ergebnisse bestätigt werden. Vor dem Hintergrund der Belastung durch das Tragen zusätzlicher Sensorik wurde für die weiteren Lernsysteme die Erkennung emotionaler Erregung ausschließlich basierend auf dem EKG-Signal vorgenommen.

Im Rahmen der Studien zeigte sich zudem, dass EDA zur Klassifikation langandauernder Erregungszustände nicht geeignet ist. In Hinblick auf die Erkennung kurzfristiger Zustandsänderungen jedoch kann das EDA-Signal eine nützliche Informationsquelle darstellen. Die Erkennung kurzfristiger Zustandsänderungen kann beim Lernen insbesondere zur Identifikation kurzfristiger Störungen verwendet werden.

Basierend auf den Studienergebnissen wurden die Informationen über das emotionale Erregungsniveau in den Lernprozess integriert. Hierbei wurden zwei verschiedene Ansätze verfolgt: Zum einen wurde ein System entwickelt, welches sich an den Nutzerzustand anpasst. Auf der anderen Seite wurde eine Trainingsanwendung vorgestellt, mittels welcher der Nutzer lernen kann, sein emotionales Erregungsniveau derart anzupassen, dass es für den Lernprozess zuträglich ist. Diese Ansätze unterscheiden sich von bisherigen Ansätzen dahingehend, dass es vielmehr um das Erlernen von Fertigkeiten geht, welche den Lernprozess unterstützen können statt um das Erlernen von Domänenwissen.

Im Rahmen der nutzerseitigen Anpassung an den Lernprozess wurde eine Studie durchgeführt, bei welcher die Nutzer über drei Wochen hinweg an einem Training der Emotionsregulationsfähigkeiten teilnahmen. Auch bei dieser Studie kamen zwei unterschiedliche Varianten des Trainings zum Einsatz: eine Variante mit und eine ohne physiologischem Feedback. Es zeigte sich, dass die Versuchspersonen, welche das Training mit physiologischem Feedback durchlaufen hatten, auch drei Wochen nach der Studie angaben, signifikant häufiger Emotionsregulationsstrategien zu verwenden, als die Versuchspersonen der Kontrollgruppe.

Zur Untersuchung der Systemadaption an den Nutzerzustand wurde eine Anwendung zum Training des Arbeitsgedächtnisses entwickelt. Diese

ist in der Lage, basierend auf Informationen aus dem EKG-Signal die Schwierigkeit des Trainings individuell an die Nutzerbedürfnisse anzupassen. Über- und Unterforderung des Nutzers können somit erkannt und die Trainingsschwierigkeit entsprechend angepasst werden. Im Rahmen einer Studie zeigte sich, dass Versuchspersonen, bei welchen die Adaption der Schwierigkeit basierend auf physiologischen Daten erfolgte, stärkere Leistungszuwächse zu verzeichnen hatten, als die Versuchspersonen einer Kontrollgruppe, bei welcher eine initiale Adaption basierend auf der Performance erfolgte.

## 9.2 Ausblick

Diese Arbeit präsentierte verschiedene Ansätze, wie physiologische Signale in Lernumgebungen integriert werden können. Sowohl hinsichtlich der Nutzeradaption im Rahmen des Emotionsregulationstrainings als auch in Bezug auf die Systemadaption beim Arbeitsgedächtnistraining zeigten sich deutliche Unterschiede zwischen der Kontrollgruppe ohne physiologischem Feedback sowie der Versuchsgruppe mit physiologischem Feedback. In zukünftigen Studien ist eine weitere Optimierung der für das physiologische Feedback verwendeten Algorithmen denkbar. Auch könnte im Rahmen einer Langzeitstudie eine stärkere Anpassung an einzelne Nutzer vorgenommen werden.

Ein Schwerpunkt dieser Arbeit lag auf der Nutzbarkeit der entwickelten Systeme unter Alltagsbedingungen. Dies wurde auch bei der Auswahl der verwendeten physiologischen Signale berücksichtigt. Schon jetzt ist eine Vielzahl an Sensoren verfügbar, welche eine nahezu belastungsfreie Messung von Vitaldaten ermöglicht. Neue Ansätze zur Bestimmung der Herzaktivität über Handykameras oder kapazitive Messungen ermöglichen bereits Aufnahmen ohne dass Sensorik direkt am Körper getragen werden muss. Zudem könnte eine Integration der entwickelten Systeme auf mobile Endgeräte erfolgen, was das Training zu jeder Zeit und an jedem Ort ermöglichen würde. In derartigen Szenarien ist eine Berücksichtigung des emotionalen Erregungsniveaus eines Nutzers als umso wichtiger einzustufen, da die Nutzer aufgrund der höheren Mobilität ständig wechselnden Umgebungseinflüssen ausgesetzt sind, welche ihr emotionales Erregungsniveau und somit auch ihre Lernleistung beeinflussen können.

Die beiden vorgestellten Systeme zum Training der Emotionsregulationsfähigkeiten sowie des Arbeitsgedächtnisses können nicht nur hinsichtlich der Verbesserung des Lernverhaltens einer Person von Vorteil sein. Auch in anderen Szenarien kann eine gezielte Steuerung des emotionalen Zustands und der Aufmerksamkeit von Nutzen sein. Interessant sind hierbei alle Bereiche, in denen schnelle und akkurate Entscheidungen getroffen werden müssen, wie es beispielsweise an Finanzmärkten der Fall ist. Eine Übertragung der Ergebnisse dieser Arbeit auf andere Anwendungsfelder ist somit gut vorstellbar.

# Literaturverzeichnis

- Alepis, E., Virvou, M. & Kabassi, K. (2006). Affective student modeling based on microphone and keyboard user actions. In *Advanced learning technologies, 2006. sixth international conference on* (S. 139–141).
- Alepis, E., Virvou, M. & Kabassi, K. (2007). Development process of an affective bi-modal intelligent tutoring system. *Intelligent Decision Technologies, 1* (3), 117–126.
- Alexander, S., Sarrafzadeh, A. & Hill, S. (2006). Easy with eve: A functional affective tutoring system. In *Workshop on Motivational and Affective Issues in ITS. 8th International Conference on ITS* (S. 5–12).
- Antonacopoulou, E. P. & Gabriel, Y. (2001). Emotion, learning and organizational change: towards an integration of psychoanalytic and other perspectives. *Journal of Organizational Change Management, 14* (5), 435–451.
- Appelhans, B. M. & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology, 10* (3), 229–240.
- Aspinwall, L. G. & Taylor, S. E. (1997). A stitch in time: self-regulation and proactive coping. *Psychological Bulletin, 121* (3), 417–436.
- Astor, P. J., Adam, M. T. P., Jerčić, P., Schaaff, K. & Weinhardt, C. (2014). Integrating biosignals into information systems: A neurois tool for improving emotion regulation. *Journal of Management Information Systems, 1*.
- Baddeley, A. (1992). Working memory. *Science, 255* (5044), 556–559.
- Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29.

- Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T. & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies*, 68 (4), 223–241.
- Bangor, A., Kortum, P. T. & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *Intl. Journal of Human-Computer Interaction*, 24 (6), 574–594.
- Becker, K. (2009). *varioport-e*. <http://www.becker-meditec.de>.
- Ben Ammar, M., Neji, M., Alimi, A. M. & Gouardères, G. (2010). The affective tutoring system. *Expert Systems with Applications*, 37 (4), 3013–3023.
- Benedek, M. & Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47 (4), 647–658.
- Benovoy, M., Cooperstock, J. R. & Deitcher, J. (2008). Biosignals analysis and its application in a performance setting. In *Proceedings of the international conference on bio-inspired systems and signal processing* (S. 253–258).
- Berger, R. D., Akselrod, S., Gordon, D. & Cohen, R. J. (1989). An efficient algorithm for spectral analysis of heart rate variability. *IEEE Transactions on Biomedical Engineering*, 33 (9), 900–904.
- Berntson, G. G., Quigley, K. S. & Lozano, D. (2007). Cardiovascular psychophysiology. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology, third edition* (S. 182–210). Cambridge: Cambridge University Press.
- Boll, H. P. (2007). *Mobile Überwachung des kardiovaskulären Systems – Konzepte und Simulationen* (Dissertation). Institut für Technik der Informationsverarbeitung, Universität Karlsruhe (TH).
- Borella, E., Carretti, B., Riboldi, F. & De Beni, R. (2010). Working memory training in older adults: evidence of transfer and maintenance effects. *Psychology and aging*, 25 (4), 767.

- Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. vollständig überarbeitete und erweiterte Auflage). Berlin: Springer.
- Bouchet, J., Nigay, L. & Ganille, T. (2004). Icare software components for rapidly developing multimodal interfaces. In *Proceedings of the 6th international conference on multimodal interfaces* (S. 251–258). New York, NY, USA: ACM.
- Boucsein, W. (2011). *Electrodermal Activity*. Springer. Hardcover.
- Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25, 49–59.
- Bradley, M. M. & Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37, 204–215.
- Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59 (1), 119–155.
- Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189, 194.
- Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121–167.
- Buschkuehl, M., Jaeggi, S. M., Hutchison, S., Perrig-Chiello, P., Däpp, C., Müller, M., . . . Perrig, W. J. (2008). Impact of working memory training on memory performance in old-old adults. *Psychology and aging*, 23 (4), 743.
- Carlson, N. R. (2007). *Physiology of behavior*. Allyn & Bacon.
- Castellano, G., Kessous, L. & Caridakis, G. (2008). Emotion recognition through multiple modalities: face, body gesture, speech. In *Affect and emotion in human-computer interaction* (S. 92–103). Springer.
- Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, 40 (3), 153–193.

- Chan, T.-W., Roschelle, J., Hsi, S., Sharples, M., Brown, T., Patton, C., ... others (2006). One-to-one technology-enhanced learning:: An opportunity for global research collaboration. *Research & Practice in Technology Enhanced Learning*, 1 (1).
- Choi, A. & Woo, W. (2005). Physiological sensing and feature extraction for emotion recognition by exploiting acupuncture spots. In *Affective computing and intelligent interaction* (S. 590–597). Springer.
- Christie, I. C. & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. *International journal of psychophysiology*, 51 (2), 143–153.
- Clauss, W. (2009). *Humanbiologie kompakt* (C. Clauss, Hrsg.). Heidelberg: Spektrum Akademischer Verlag.
- Clifford, G. D. (2007). ECG statistics, noise, artifacts, and missing data. In R. Acharya, J. Suri, J. Spaan & S. Krishnan (Hrsg.), *Advanced methods and tools for ECG data analysis* (S. 55–100). Artech House, Boston.
- Coan, J. A. & Allen, J. J. (2007). *Handbook of emotion elicitation and assessment*. Oxford university press.
- Cogmed. (2014). *Cogmed solutions help consumers, professionals, and schools address attention problems*. Zugriff auf <http://www.cogmed.com/program> (abgerufen am 03.06.2014)
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O. & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12 (5), 769–786.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety: the experience of play in work and games*. San Francisco: Jossey-Bass.
- Csikszentmihalyi, M. & Csikszentmihalyi, I. S. (1988). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge University Press.
- Curtin, J. J., Lozano, D. L. & Allen, J. J. (2007). The psychophysiological laboratory. *Handbook of emotion elicitation and assessment*, 398–425.

- Cuthbert, B. N., Bradley, M. M. & Lang, P. J. (1996). Probing picture perception: activation and emotion. *Psychophysiology*, 33 (2), 103–111.
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in w/m and reading, vol. 19, pp. 450-466. *Journal of Verbal Learning and Verbal Behavior*, 19, 450–466.
- Daneman, M. & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3 (4), 422–433.
- Dawson, M. E., Schell, A. M. & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Hrsg.), *Handbook of psychophysiology, third edition* (S. 159–181). Cambridge: Cambridge University Press.
- Degen, R. (2011). *Entwicklung einer drucksensitiven Maus für die Durchführung physio-ökonomischer Studien* (Diplomarbeit). FZI Forschungszentrum Informatik, Karlsruher Institut für Technologie (KIT).
- Dennerlein, J., Becker, T., Johnson, P., Reynolds, C. J. & Picard, R. W. (2003, 24.). Frustrating Computer Users Increases Exposure to Physical Factors. In *Proceedings of the international ergonomics association*. Seoul, Korea.
- Diamond, D. M., Campbell, A. M., Park, C. R., Halonen, J. & Zoladz, P. R. (2007). The temporal dynamics model of emotional memory processing: a synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the yerkes-dodson law. *Neural Plasticity*, 2007, 1–33.
- Dickman, S. J. (2002). Dimensions of arousal: wakefulness and vigor. *Human Factors*, 44 (3), 429–442.
- Dimberg, U. (1990, Sep). Facial electromyography and emotional reactions. *Psychophysiology*, 27 (5), 481–494.
- D’Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H., ... others (2008). Autotutor detects and responds to learners affective and cognitive states. In *Workshop on emotional and cognitive issues at the international conference on intelligent tutoring systems*.



- D'Mello, S. & Kory, J. (2012). Consistent but modest: A meta-analysis on unimodal and multimodal affect detection accuracies from 30 studies. In *Proceedings of the 14th ACM international conference on multimodal interaction* (S. 31–38). New York, NY, USA: ACM.
- D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., . . . Graesser, A. (2010). A time for emoting: When affect-sensitivity is and isn't effective at promoting deep learning. In V. Alevan, J. Kay & J. Mostow (Hrsg.), *Intelligent tutoring systems* (Bd. 6094, S. 245–254). Springer Berlin Heidelberg.
- D'Mello, S., Picard, R. & Graesser, A. (2007). Towards an affect-sensitive autotutor. *Special issue on Intelligent Educational Systems - IEEE Intelligent Systems*, 22 (4), 53–61.
- Donato, G., Bartlett, M. S., Hager, J. C., Ekman, P. & Sejnowski, T. J. (1999). Classifying facial actions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21 (10), 974–989.
- Dragicevic, P. & Fekete, J. (2001). Input device selection and interaction configuration with icon. *People and Computers*, 543–448.
- Dror, I. E. (2008). Technology enhanced learning: The good, the bad, and the ugly. *Pragmatics & Cognition*, 16 (2), 215–223.
- du Boulay, B., Avramides, K., Luckin, R., Martínez-Mirón, E., Méndez, G. R. & Carr, A. (2010). Towards systems that care: a conceptual framework based on motivation, metacognition and affect. *International Journal of Artificial Intelligence in Education*, 20 (3), 197–229.
- Dumas, B., Lalanne, D. & Oviatt, S. (2009). Human machine interaction. In D. Lalanne & J. Kohlas (Hrsg.), (S. 3–26). Berlin, Heidelberg: Springer-Verlag.
- Ebner-Priemer, U. W. & Kubiak, T. (2007). Psychological and psychophysiological ambulatory monitoring: A review of hardware and software solutions. *European Journal of Psychological Assessment*, 23 (4), 214–226.

- Eich, E., Ng, J. T., Macaulay, D., Percy, A. D. & Grebneva, I. (2007). Combining music with thought to change mood. *Handbook of emotion elicitation and assessment*, 124–136.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169–200.
- Ekman, P. & Friesen, W. V. (1977). *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press, Stanford University, Palo Alto.
- Elter, P. (2001). *Methoden und Systeme zur nichtinvasiven, kontinuierlichen und belastungsfreien Blutdruckmessung* (Dissertation). Institut für Technik der Informationsverarbeitung, Universität Karlsruhe (TH).
- Eyben, F., Wöllmer, M., Poitschke, T., Schuller, B., Blaschke, C., Färber, B. & Nguyen-Thien, N. (2010). Emotion on the road—necessity, acceptance, and feasibility of affective computing in the car. *Advances in human-computer interaction, 2010*.
- Eysenck, M. W. & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, 6 (6), 409–434.
- Fenton-O’Creevy, M., Conole, G., Lins, J., Peffer, G., Adam, M., Lindley, C., ... Scanlon, E. (2012). A learning design to support the emotion regulation of investors.
- Fleureau, J., Guillotel, P. & Orlac, I. (2013). Affective benchmarking of movies based on the physiological responses of a real audience. In *International conference on affective computing & intelligent interaction. (ACII 2013)* (S. 73–78).
- Fong, T., Nourbakhsh, I. & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and autonomous systems*, 42 (3), 143–166.
- Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T. & Venables, P. H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology*, 18, 232–239.
- Fraden, J. (2004). *Handbook of modern sensors: physics, designs and applications* (3. Auflage). New York, NY: Springer.

- Gay, V., Leijdekkers, P., Agcanas, J., Wong, F. & Wu, Q. (2013). Captu-remyemotion: Helping autistic children understand their emotions using facial expression recognition and mobile technologies. In *innovations: Challenges and impacts for individuals, organizations and society* (S. 409–420).
- Geddes, L. A., Voelz, M. H., Babbs, C. F., Bourland, J. D. & Tacker, W. A. (1981). Pulse transit time as an indicator of arterial blood pressure. *Psychophysiology*, 18 (1), 71–74.
- Gharbi, A., Hey, S., Jatoba, L., Grossmann, U., Ottenbacher, J., Kuncoro, C., ... Müller-Glaser, K. D. (2008). System for body and mind monitoring in coaching process. In *Medical devices and biosensors, 2008. (ISSS-MDBS 2008). 5th international summer school and symposium on* (S. 89–91).
- Gilleade, K., Dix, A. & Allanson, J. (2005). Affective videogames and modes of affective gaming: assist me, challenge me, emote me. In *Digra 2005 conference: Changing views – worlds in play, proceedings of*.
- Goedhart, A. D., van der Sluis, S., Houtveen, J. H., Willemsen, G. & de Geus, E. J. (2007). Comparison of time and frequency domain measures of rsa in ambulatory recordings. *Psychophysiology*, 44 (2), 203–215.
- Gomez, P. & Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, 53 (2), 91–103.
- Gootjes, L., Coppens, L. C., Zwaan, R. A., Franken, I. H. & Strien, J. W. V. (2011). Effects of recent word exposure on emotion-word stroop interference: An ERP study. *International Journal of Psychophysiology*, 79 (3), 356–363.
- Gramann, K. & Schandry, R. (2009). *Psychophysiologie : Körperliche Indikatoren psychischen Geschehens* (4. vollständig überarbeitete Auflage). Weinheim [u.a.]: Beltz.
- Grandjean, E. (1991). *Physiologische Arbeitsgestaltung : Leitfaden der Ergonomie* (Sonderausgabe, 4. erweiterte Auflage). Landsberg: ecomed.

- Grillon, C., Pine, D. S., Baas, J. M. P., Lawley, M., Ellis, V. & Charney, D. S. (2006). Cortisol and dhea-s are associated with startle potentiation during aversive conditioning in humans. *Psychopharmacology*, *186* (3), 434–441.
- Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74* (1), 224–237.
- Gross, J. J. & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology*, *85* (2), 348–362.
- Gross, J. J. & Levenson, R. W. (1993). Emotional suppression: physiology, self-report, and expressive behavior. *Journal of Personality and Social Psychology*, *64* (6), 970–986.
- Hale, S., Rose, N. S., Myerson, J., Strube, M. J., Sommers, M., Tye-Murray, N. & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging*, *26* (1), 92–110.
- Hamilton, P. (2002). Open source ecg analysis. In *Computers in cardiology, 2002* (S. 101–104).
- Harbison, J. I., Atkins, S. & Dougherty, M. R. (2011). N-back training task performance: Analysis and model. In *Proceedings of the cognitive science society*.
- Heinzel, S., Schulte, S., Onken, J., Duong, Q.-L., Riemer, T. G., Heinz, A., ... Rapp, M. A. (2013). Working memory training improvements and gains in non-trained cognitive tasks in young and older adults. *Aging, Neuropsychology, and Cognition*, 1–28.
- Hering, E. & Schönfelder, G. (2011). *Sensoren in Wissenschaft und Technik: Funktionsweise und Einsatzgebiete*. Springer.
- Hernandez, J., McDuff, D., Fletcher, R. & Picard, R. W. (2013). Inside-out: Reflecting on your inner state. In *Pervasive computing and communications workshops (percom workshops), 2013 IEEE international conference on* (S. 324–327).

- Hjelm, S. I. (2003). Research+ design: the making of brainball. *Interactions*, 10 (1), 26–34.
- Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29 (3), 306–310.
- Holmes, J., Gathercole, S. E. & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental science*, 12 (4), F9–F15.
- Ivonin, L., Chang, H.-M., Chen, W. & Rauterberg, M. (2013). Automatic recognition of the unconscious reactions from physiological signals. In *Human factors in computing and informatics* (S. 16–35). Springer.
- Izard, C. E. (1994). *Die Emotionen des Menschen* (3. Auflage). Beltz, Psychologie-Verl.-Union.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J. & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, 105 (19), 6829–6833.
- Jaeggi, S. M., Seewer, R., Nirkko, A. C., Eckstein, D., Schroth, G., Groner, R. & Gutbrod, K. (2003). Does excessive memory load attenuate activation in the prefrontal cortex? load-dependent processing in single and dual tasks: functional magnetic resonance imaging study. *Neuroimage*, 19 (2), 210–225.
- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y.-F., Jonides, J. & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning—implications for training and transfer. *Intelligence*, 38 (6), 625–635.
- Jerčić, P., Astor, P. J., Adam, M. T. P., Hilborn, O., Schaaff, K., Lindley, C., ... Eriksson, J. (2012). A serious game using physiological interfaces for emotion regulation training in the context of financial decision-making. In *European conference on information systems (ECIS)*.
- John, O. P. & Gross, J. J. (2004). Healthy and unhealthy emotion regulation: personality processes, individual differences, and life span development. *Journal of Personality*, 72 (6), 1301–1333.

- Jungle Memory. (2014). *Who it helps: Find success with jungle memory*. Zugriff auf [http://junglememory.com/pages/who\\_can\\_benefit](http://junglememory.com/pages/who_can_benefit) (abgerufen am 03.06.2014)
- Kahn, J., Ducharme, P., Travers, B. & Gonzalez-Heydrich, J. (2009). Rage control: Regulate and gain emotional control. *Strategy for the Future of Health*, 3, 335.
- Kim, J. & André, E. (2008). Emotion recognition based on physiological changes in music listening. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30 (12), 2067–2083.
- Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology*, 55 (4), 352–358.
- Kirschbaum, C., Pirke, K.-M. & Hellhammer, D. H. (1993). The 'trier social stress test'—a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28 (1-2), 76–81.
- Klingberg, T. (2010). Training and plasticity of working memory. *Trends in cognitive sciences*, 14 (7), 317–324.
- Klingberg, T., Forssberg, H. & Westerberg, H. (2002). Training of working memory in children with adhd. *Journal of clinical and experimental neuropsychology*, 24 (6), 781–791.
- Kroschel, K., Rigoll, G., Schuller, B. & Kroschel, K. (2011). *Statistische Informationstechnik : Signal - und Mustererkennung, Parameter- und Signalschätzung*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Lane, J. D., Greenstadt, L., Shapiro, D. & Rubinstein, E. (1983). Pulse transit time and blood pressure: an intensive analysis. *Psychophysiology*, 20 (1), 45–49.
- Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. Sidowski, J. Johnson & T. Williams (Hrsg.), *Technology in mental health care delivery systems* (S. 119–139). Norwood, NJ: Ablex.

- Lang, P. J., Bradley, M. M. & Cuthbert, B. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual* (Bericht Nr. A-6). Gainesville, FL: University of Florida.
- Lang, P. J., Greenwald, M. K., Bradley, M. M. & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. In (Bd. 30, S. 261–273).
- Lawson, J.-Y. L., Vanderdonckt, J. & Macq, B. (2010). OpenInterface: A lightweight open source platform for rapid prototyping of multimodal applications. In *Proceedings of 1st international workshop on user interface extensible markup language (UsiXML2010)*.
- Litman, D. J. & Silliman, S. (2004). ITSPOKE: An intelligent tutoring spoken dialogue system. In *Demonstration papers at HLT-NAACL 2004* (S. 5–8).
- Liu, C., Agrawal, P., Sarkar, N. & Chen, S. (2009). Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback. *International Journal of Human-Computer Interaction*, 25 (6), 506–529.
- Lohaus, A. & Vierhaus, M. (2013). Intelligenz. In *Entwicklungspsychologie des Kindes- und Jugendalters für Bachelor* (S. 122–136). Springer Berlin Heidelberg.
- Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39 (2), 447–462.
- Lozano, L. M., García-Cueto, E. & Muñoz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 4 (2), 73–79.
- MacLeod, C. M. (1991). Half a century of research on the stroop effect: an integrative review. *Psychological Bulletin*, 109 (2), 163–203.
- Malik, M. (1996). Heart rate variability : Standards of measurement, physiological interpretation, and clinical use. task force of the european society of cardiology the north american society of pacing and electrophysiology. *Circulation*, 93 (5), 1043–1065.

- Mandryk, R. L. & Atkins, M. S. (2007). A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 65 (4), 329–347.
- Mathews, A. & MacLeod, C. (1985). Selective processing of threat cues in anxiety states. *Behaviour Research and Therapy*, 23 (5), 563–569.
- Mietus, J., Peng, C.-K. & Henry, I. (2002). The pNNx-files: Reexamining a widely used heart rate variability measure. *Heart*, 88, 378–380.
- Mindware Consulting, Inc. (2014). *Mind games*. Zugriff auf <https://play.google.com/store/apps/details?id=mindware.mindgames> (abgerufen am 03.06.2014)
- Moody, G. B. (2013). *WFDB programmer's guide* (Bericht Nr. Tenth Edition (revised and with additions for WFDB library version 10.5.18), 22 February 2013). Harvard-MIT Division of Health Sciences and Technology.
- movisens GmbH. (2014). *ekgMove - Mobile ECG and Physical Activity Sensor*.
- Nakamura, J. & Csikszentmihalyi, M. (2002). Handbook of positive psychology. In C. R. Snyder & S. J. Lopez (Hrsg.), (S. 89–105). Oxford University Press.
- NeuroNation. (2014a). *Das Erfolgstraining für Ihren Kopf*. Zugriff auf <http://www.neuronation.de> (abgerufen am 03.06.2014)
- NeuroNation. (2014b). *NeuroNation – Gehirn Training*. Zugriff auf <https://play.google.com/store/apps/details?id=air.nn.mobile.app.main> (abgerufen am 03.06.2014)
- Newlin, D. B. (1981). Relationships of pulse transmission times to pre-ejection period and blood pressure. *Psychophysiology*, 18 (3), 316–321.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, 84 (3), 231.
- Nkambou, R. (2006). Managing student emotions in intelligent tutoring systems. In *FLAIRS conference* (S. 389–394).



- Novak, D., Mihelj, M. & Munih, M. (2012). A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interacting with Computers*, 24 (3), 154–172.
- Nussinovitch, U., Elishkevitz, K. P., Katz, K., Nussinovitch, M., Segev, S., Volovitz, B. & Nussinovitch, N. (2011). Reliability of ultra-short ECG indices for heart rate variability. *Annals of Noninvasive Electrocardiology*, 16 (2), 117–122.
- Öhman, A., Flykt, A. & Esteves, F. (2001). Emotion drives attention: detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130 (3), 466–478.
- Oliveira, E., Benovoy, M., Ribeiro, N. & Chambel, T. (2011). Towards emotional interaction: using movies to automatically learn users' emotional states. In *Human-computer interaction-INTERACT 2011* (S. 152–161). Springer.
- Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49, 197–237.
- Ottenbacher, J., Kirst, M., Jatobá, L., Huflejt, M., Großmann, U. & Stork, W. (2008). Reliable motion artifact detection for ecg monitoring systems with dry electrodes. In *Engineering in medicine and biology society, 2008. EMBS 2008. 30th annual international conference of the IEEE* (S. 1695–1698).
- Pantelopoulos, A. & Bourbakis, N. G. (2010). A survey on wearable sensor-based systems for health monitoring and prognosis. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40 (1), 1–12.
- Park, D. C. & Payer, D. (2006). Working memory across the adult lifespan. In E. Bialystock & F. I. M. Craik (Hrsg.), *Life span cognition: Mechanisms of change* (S. 128–142). New York, NY: Oxford University Press.
- Patel, S., Park, H., Bonato, P., Chan, L. & Rodgers, M. (2012). A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation*, 9 (1), 21.

- Payne, R. A., Symeonides, C. N., Webb, D. J. & Maxwell, S. R. J. (2006). Pulse transit time measured from the ECG: an unreliable marker of beat-to-beat blood pressure. *Journal of Applied Physiology*, 100 (1), 136–141.
- Pearson Education, Inc. Mobile. (2013). *Cogmed*. Zugriff auf <https://play.google.com/store/apps/details?id=air.com.pearsoned.cogmed.rm> (abgerufen am 03.06.2014)
- Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational psychologist*, 37 (2), 91–105.
- Picard, R. (2009). Future effective technology for autism and emotion communication. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 364 (1535), 3575–3584.
- Picard, R. W. (1995). *Affective computing* (Bericht Nr. 321). Cambridge, MA: MIT Media Laboratory, Perceptual Computing Section.
- Pollack, I., Johnson, L. B. & Robert, K. P. (1959). Running memory span. *Journal of Experimental Psychology*, 57 (3), 137–146.
- Rani, P., Liu, C., Sarkar, N. & Vanman, E. (2006). An empirical study of machine learning techniques for affect recognition in human–robot interaction. *Pattern Analysis and Applications*, 9 (1), 58–69.
- Reynolds, C. & Reynolds, C. J. (2001). *The sensing and measurement of frustration with computers*.
- Richmond, L. L., Morrison, A. B., Chein, J. M. & Olson, I. R. (2011). Working memory training and transfer in older adults. *Psychology and aging*, 26 (4), 813.
- Rickles, W. H. J. & Day, J. (1968). Electrodermal activity in non-palmar skin sites. *Psychophysiology*, 4 (4), 421–435.
- Ritterfeld, U., Cody, M. & Vorderer, P. (2009). *Serious games: Mechanisms and effects*. Routledge.

- Robinson, M. D. & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological bulletin*, 128 (6), 934.
- Russel, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology*, 37, 345–356.
- Russel, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39, 1161–1178.
- Russel, J. A. & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11, 273–294.
- Russell, J. & Fehr, B. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113, 464–486.
- Sakr, G. E., Elhajj, I. H. & Huijjer, H.-S. (2010). Support vector machines to define and detect agitation transition. *Affective Computing, IEEE Transactions on*, 1 (2), 98–108.
- Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C. & Bigdeli, A. (2008). 'How do you know that I don't understand?' A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24 (4), 1342–1363.
- Scargle, J. D. (1982). Study in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263, 835–853.
- Schaaff, K. (2013). Enhancing mobile working memory training by using affective feedback. In *IADIS International Conference on Mobile Learning, 14-16 March 2013, Lisbon, Portugal*.
- Schaaff, K. & Adam, M. T. P. (2013). Measuring emotional arousal for online applications: Evaluation of ultra-short term heart rate variability measures. In *International conference on affective computing & intelligent interaction. (ACII 2013)*.
- Schaaff, K., Degen, R., Adler, N. & Adam, M. T. (2012). Measuring affect using a standard mouse device. *Biomedizinische Technik*, 57, 1.

- Schaaff, K., Müller, L., Kirst, M. & Heuer, S. (2012). xAffect - a modular framework for online affect recognition and biofeedback applications. In *7th European Conference on Technology Enhanced Learning (ECTEL 2012), MATEL Workshop, Saarbrücken, Germany*.
- Schaaff, K. & Schultz, T. (2009). Towards an EEG-based emotion recognizer for humanoid robots. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on* (S. 792–796).
- Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N. & Wolpaw, J. R. (2004). BCI2000: a general-purpose brain-computer interface (BCI) system. *Biomedical Engineering, IEEE Transactions on*, 51 (6), 1034–1043.
- Schandry, R. (1989). *Lehrbuch der Psychophysiologie* (2., überarbeitete und erweiterte Auflage). Psychologie Verlags Union.
- Schestatsky, P., Callejas, M. A. & Valls-Sole, J. (2010). Abnormal modulation of electrodermal activity by thermoalgesic stimuli in patients with primary palmar hyperhidrosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 82, 92–96.
- Schmidt-Atzert, L. (1981). *Emotionspsychologie*. Kohlhammer.
- Schmiedek, F., Lövdén, M. & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the cogito study. *Frontiers in Aging Neuroscience*, 2, 1–10.
- Schneider, J., Koellner, C. & Heuer, S. (2012). An approach to automotive ECG measurement validation using a car-integrated test framework. In *Intelligent Vehicles Symposium (IV), 2012 IEEE* (S. 950–955).
- Schroeder, E. B., Whitsel, E. A., Evans, G. W., Prineas, R. J., Chambless, L. E. & Heiss, G. (2004). Repeatability of heart rate variability measures. *Journal of Electrocardiology*, 37 (3), 163–172.
- Sharry, J., McDermott, M. & Condrón, J. (2003). Relax to win treating children with anxiety problems with a biofeedback video game. *Eisteach*, 2, 22–26.

- Shipstead, Z., Redick, T. S. & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, 138 (4), 628–654.
- Silbernagl, S. & Despopoulos, A. (2007). *Taschenatlas Physiologie* (7. Auflage). Thieme.
- Singh, D., Vinod, K. & Saxena, S. C. (2004). Sampling frequency of the rr interval time series for spectral analysis of heart rate variability. *Journal of Medical Engineering & Technology*, 28 (6), 263–272.
- Sinha, A. K. & Landay, J. A. (2003). Capturing user tests in a multimodal, multidevice informal prototyping tool. In *Proceedings of the 5th international conference on multimodal interfaces* (S. 117–124). New York, NY, USA: ACM.
- Spencer, H. (1890). *The principles of psychology* (Bd. 1). New York: Appleton.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of experimental psychology*, 18 (6), 643–662.
- Sykes, J. & Brown, S. (2003). Affective gaming: Measuring emotion through the gamepad. In *Chi '03 extended abstracts on human factors in computing systems* (S. 732–733). New York, NY, USA: ACM.
- Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G. & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental science*, 12 (1), 106–113.
- Tobola, A. & Douniama, C. (2007). Evaluation of alternative derivation areas for plethysmography and pulse oximetry. In *Sensation 2nd international conference, monitoring sleep and sleepiness with new sensors within medical and industrial applications*.
- Tullis, T. S. & Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. In *Usability professional association conference* (S. 1–12).
- Tulppo, M., Makikallio, T., Takala, T., Seppanen, T. & Huikuri, H. (1996). Quantitative beat-to-beat analysis of heart rate dynamics during exercise. *American Journal of Physiology*, 271, H244–H252.

- Urbian. (2014). *Memory trainer*. Zugriff auf <https://play.google.com/store/apps/details?id=org.urbian.android.games.memorytrainer> (abgerufen am 03.06.2014)
- Valenza, G., Lanata, A. & Scilingo, E. P. (2012). The role of nonlinear dynamics in affective valence and arousal recognition. *Affective Computing, IEEE Transactions On*, 3 (2), 237–249.
- van den Broek, E. L., Janssen, J. H., Healey, J. A. & van der, M. Z. (2010, January). Prerequisites for affective signal processing (ASP) – part ii. In A. Fred, J. Filipe & H. Gamboa (Hrsg.), *Biosignals 2010: proceedings of the third international conference on bio-inspired systems and signal processing* (S. 188–193). Portugal: INSTICC Press.
- van den Broek, E. L., Janssen, J. H., van der, M. D. Z., Westerink, J. H. & Healey, J. A. (2011, January). Prerequisites for affective signal processing (ASP) – part v: A response to comments and suggestions. In F. Babiloni, A. Fred, J. Filipe & H. Gamboa (Hrsg.), *Third international conference on bio-inspired systems and signal processing, biosignals 2011* (S. 301–306). INSTICC Press.
- van den Broek, E. L., Janssen, J. H., Westerink, J. H. & Healey, J. A. (2009). Prerequisites for affective signal processing (ASP). In P. Encarnacao & A. Veloso (Hrsg.), *International conference on bio-inspired systems and signal processing, biosignals 2009* (S. 426–433). Portugal: INSTICC Press.
- Vuilleumier, P. & Schwartz, S. (2001). Beware and be aware: capture of spatial attention by fear-related stimuli in neglect. *Neuroreport*, 12 (6), 1119–1122.
- Wallin, B. G. (1981). Sympathetic nerve activity underlying electrodermal and cardiovascular reactions in man. *Psychophysiology*, 18 (4), 470–476.
- Watson, D., Clark, L. A. & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54 (6), 1063.
- Woolf, B., Dragon, T., Arroyo, I., Cooper, D., Bursleson, W. & Muldner, K. (2009). Recognizing and responding to student affect. In J. Jacko

- (Hrsg.), *Human-computer interaction. ambient, ubiquitous and intelligent interaction* (Bd. 5612, S. 713–722). Springer Berlin / Heidelberg.
- Wundt, W. M. (1896). *Grundriss der Psychologie*. Leipzig: Wilhelm Engelmann.
- Yerkes, R. & Dodson, J. (1908). The relation of strength of stimulus to the rapidity of habit formation. *Journal of Comparative Neurology and Psychology*, 18, 459–482.
- Zeelenberg, M. & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, 17 (1), 3–18.
- Zeng, Z., Pantic, M., Roisman, G. I. & Huang, T. S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31 (1), 39–58.
- Zhou, F., Qu, X., Helander, M. G. & Jiao, J. R. (2011). Affect prediction from physiological measures via visual stimuli. *International Journal of Human-Computer Studies*, 69 (12), 801–819.
- Zimmermann, C. (2012). *Entwicklung und Evaluation von Algorithmen für die Online- und Offline-Analyse von EDA-Signalen*. Karlsruher Institut für Technologie (KIT), Diplomarbeit.

# Eigene Veröffentlichungen

- Astor, P. J., Adam, M. T. P., Jerčić, P., Schaaff, K. & Weinhardt, C. (2014). Integrating biosignals into information systems: A neurois tool for improving emotion regulation. *Journal of Management Information Systems*, 1.
- Fenton-O’Creevy, M., Lins, J. T., Vohra, S., Richards, D. W., Davies, G. & Schaaff, K. (2012a). Emotion regulation and trader expertise: Heart rate variability on the trading floor. *Journal of Neuroscience, Psychology, and Economics*, 5 (4), 227–237.
- Fenton-O’Creevy, M., Lins, J. T., Vohra, S., Richards, D. W., Davies, G. & Schaaff, K. (2012b). Putting emotion at the heart of finance: Emotion regulation and trader expertise. In *Cass Business School Behavioural Finance Conference, London, UK, 21-22 June 2012*.
- Jerčić, P., Astor, P. J., Adam, M. T. P., Hilborn, O., Schaaff, K., Lindley, C. A., ... Eriksson, J. (2012). A serious game using physiological interfaces for emotion regulation training in the context of financial decision-making. *ECIS 2012 Proceedings*, 20.
- Kapp, D., Schaaff, K., Ottenbacher, J., Heuer, S. & Stork, W. (2014). Evaluation of environmental effects on the measurement of electrodermal activity under real-life conditions. *Biomedizinische Technik*.
- Kunze, C., Schaaff, K. & Hey, S. (2009). Sensor technologies for psychophysiological monitoring and ambulatory assessment. In *1st Intl. Workshop on Physioeconomics (PHYSIOECONOMICS 2009), Rotterdam, 2009*.
- Schaaff, K. (2011). Biosignalbasiertes Training des Emotionserlebens. In *MinD-Akademie 2011 – Zukunft und Forschung, Hannover, 2011*.



- Schaaff, K. (2013). Enhancing mobile working memory training by using affective feedback. In *IADIS International Conference on Mobile Learning, 14-16 March 2013, Lisbon, Portugal*.
- Schaaff, K. & Adam, M. T. P. (2013). Measuring emotional arousal for online applications: Evaluation of ultra-short term heart rate variability measures. In *International conference on affective computing & intelligent interaction. (ACII 2013)*.
- Schaaff, K., Degen, R., Adler, N. & Adam, M. T. (2012). Measuring affect using a standard mouse device. *Biomedizinische Technik*, 57, 1.
- Schaaff, K., Kirst, M. & Müller, L. (2011). Ein psychophysiologisches Monitoring-Framework zur Verbesserung der Entscheidungskompetenz an Finanzmärkten. In *7. Tagung der Fachgruppe Arbeits-, Organisations- und Wirtschaftspsychologie der Deutschen Gesellschaft für Psychologie*.
- Schaaff, K., Müller, L., Kirst, M. & Heuer, S. (2011). Improving financial decision making using psychophysiological sensor data. In *51st annual meeting of the society of psychophysiological research spr*.
- Schaaff, K., Müller, L., Kirst, M. & Heuer, S. (2012). xAffect - a modular framework for online affect recognition and biofeedback applications. In *7th European Conference on Technology Enhanced Learning (ECTEL 2012), MATEL Workshop, Saarbrücken, Germany*.
- Schaaff, K. & Schultz, T. (2009a). EEG-based emotion recognition using support vector machines. In *1. Fachtagung Biophysiological Interfaces, Berlin, Germany, 12.06.2009*.
- Schaaff, K. & Schultz, T. (2009b). Towards an EEG-based emotion recognizer for humanoid robots. In *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on* (S. 792–796).
- Schaaff, K. & Schultz, T. (2009c). Towards emotion recognition from electroencephalographic signals. In *3rd international conference on affective computing and intelligent interaction, 2009. ACII 2009* (S. 1–6).

# Betreute studentische Arbeiten

- Degen, R. (2011). *Entwicklung einer drucksensitiven Maus für die Durchführung physio-ökonomischer Studien*. Karlsruher Institut für Technologie (KIT), Diplomarbeit.
- Hettmann, H. (2014). *Bestimmung emotionaler Erregung aus multimodalen Sensordaten*. Karlsruher Institut für Technologie (KIT), Masterarbeit.
- Kapp, D. M. (2014). *Bewertung von Umgebungseinflüssen auf die Messung von elektrodermaler Aktivität unter Alltagsbedingungen*. Karlsruher Institut für Technologie (KIT), Bachelorarbeit.
- Maurer, T. (2011). *Elektrodermale Aktivität - Signalverarbeitung und Analyse*. Karlsruher Institut für Technologie (KIT), Bachelorarbeit.
- Mokin, A. (2013). *Entwicklung eines biosignalbasierten Verfahrens zur Erkennung emotionaler Erregung*. Karlsruher Institut für Technologie (KIT), Masterarbeit.
- Weiss, M. (2014). *Entwicklung eines adaptiven Systems zum Training des Arbeitsgedächtnisses basierend auf physiologischen Signalen*. Karlsruher Institut für Technologie (KIT), Masterarbeit.
- Zimmermann, C. (2012). *Entwicklung und Evaluation von Algorithmen für die Online- und Offline-Analyse von EDA-Signalen*. Karlsruher Institut für Technologie (KIT), Diplomarbeit.



# Abbildungsverzeichnis

1.1	Regelkreis des technologiegestützten Lernens . . . . .	3
1.2	Verknüpfte Forschungsgebiete im Rahmen dieser Arbeit .	4
2.1	Zusammenwirken von sympathischen und parasympathischen Nervensystem . . . . .	9
2.2	Schematische Darstellung des Herz-Kreislaufsystems . . .	10
2.3	EKG-Signal in Abhängigkeit von der Ableitposition . . . .	11
2.4	Typischer Verlauf eines EKG-Signals . . . . .	12
2.5	Arten von PPG-Messungen: A transmissiv, B reflexiv . . .	13
2.6	Schematische Darstellung einer Pulswelle . . . . .	14
2.7	Anatomie einer ekkrinen Schweißdrüse in den unterschiedlichen Hautschichten . . . . .	16
2.8	Standard-Positionen für EDA-Ableitungen an der Handfläche mit Elektrodenplatzierungen . . . . .	18
2.9	Aufbau eines FSRs . . . . .	19
2.10	Exemplarische Darstellung des zweidimensionalen Emotionsmodells . . . . .	22
2.11	Zusammenhang von Leistung und emotionaler Erregung nach dem Yerkes-Dodson-Gesetz . . . . .	24
2.12	Das Flow-Modell nach Csikszentmihalyi . . . . .	24
2.13	Ablauf der $n$ -Back-Task . . . . .	28
2.14	Prozessmodell der Emotionsregulation . . . . .	29
2.15	Ermittlung der Herzrate aus dem EKG-Signal . . . . .	32
2.16	Charakteristische Merkmale der Pulswelle . . . . .	35
2.17	Berechnung der PTT als Differenz zwischen R-Zacke und dem virtuellen Fußpunkt . . . . .	36
2.18	Trennung des EDA-Signals in phasische und tonische Aktivität	37
2.19	Typische Parameter elektrodermalen Reaktionen . . . . .	38
2.20	Vorgehensweise bei der Detektion elektrodermalen Reaktionen	38

3.1	Unterschiedliche Arten mobiler Polygraphen zur Aufzeichnung physiologischer Signale . . . . .	49
3.2	SAM-Fragebogen für Valenz und Erregung . . . . .	51
3.3	Schritte bei der Emotionsklassifikation . . . . .	56
4.1	Aufbau des Gesamtsystems . . . . .	68
4.2	Mind and Body Monitoring Labormesssystem . . . . .	72
4.3	ekgMove-Sensor . . . . .	73
4.4	Integration der Sensorik in die Maus . . . . .	77
4.5	Gesamtarchitektur der xAffect-Software . . . . .	79
4.6	Elementare Datenflussmuster . . . . .	82
4.7	Lebenszyklus der Komponenten . . . . .	83
4.8	Nutzeroberfläche der xAffect-Software . . . . .	86
4.9	Varianten zur Integration von physiologischem Feedback in Lernanwendungen . . . . .	90
5.1	Spieldesign zur Induktion hoher emotionaler Erregung . .	93
5.2	Spieldesign zur Induktion niedriger emotionaler Erregung	94
5.3	Setup der xAffect-Software für die Studiendurchführung .	95
5.4	Klassifikationskonzept . . . . .	97
5.5	Relative Veränderung der EKG-basierten Features im Vergleich zur Baseline . . . . .	108
5.6	Klassifikationsgenauigkeiten verschiedener Klassifikatoren für emotionale Erregung auf EKG-basierten Features . . .	109
5.7	Relative Veränderung von PPG-Amplitude und PTT im Vergleich zur Baseline gemittelt nach experimenteller Bedingung . . . . .	111
5.8	Relative Veränderung von Amplitude und Anzahl der Nulldurchgänge während des Spiels im Vergleich zur Baseline	113
5.9	Klassifikationsgenauigkeiten verschiedener Klassifikatoren für emotionale Erregung basierend auf dem Zeitverlauf des EDA-Signals . . . . .	114
5.10	Streuung der Klickkraft während des Spiels im Vergleich zur Baseline für beide experimentellen Bedingungen . . .	116
5.11	Relative Veränderung der Mausklicks während des Spiels .	117
5.12	Klassifikationsgenauigkeiten für emotionale Erregung basierend auf den mausbasierten Features . . . . .	118

---

5.13	Zeitlicher Ablauf der Studie und Einteilung der Phasen in unterschiedliche Erregungsniveaus . . . . .	121
5.14	Mittlerer Cortisolspiegel über alle Versuchspersonen für die einzelnen Phasen des Experiments . . . . .	122
5.15	Mittlere MDBF-Werte über alle Versuchspersonen für die einzelnen Phasen des Experiments . . . . .	123
5.16	Einteilung der experimentellen Phasen in unterschiedliche Erregungsstufen . . . . .	124
5.17	Gesamtklassifikationsergebnisse bei zwei Klassen für die einzelnen physiologischen Signale sowie deren Kombination	125
5.18	Klassifikationsergebnisse bei vier Klassen für die einzelnen physiologischen Signale sowie deren Kombination . . . . .	125
6.1	Zeitlicher Ablauf der Studie. Rote Zahlen geben die Anzahl der emotionalen Stimuli in jeder Phase an . . . . .	131
6.2	Sensitivität verschiedener Methoden zur Detektion von EDA-Reaktionen . . . . .	134
6.3	Vergleich der mittleren Amplitudenhöhe emotionaler und nicht-emotionaler EDA-Reaktionen . . . . .	135
6.4	F-Wert für die Detektion emotionaler EDA-Reaktionen in Abhängigkeit des Schwellwertes für die Amplitudenhöhe .	136
6.5	Anteil der akzeptierten EDA-Daten in Abhängigkeit von Aktivität und Herzrate . . . . .	137
6.6	Normierter Mittelwert über die physiologischen Parameter aller Versuchspersonen. Rote Linien kennzeichnen emotionale Stimuli während der Durchführung des Experiments	140
7.1	Spielaufbau des Auction Games . . . . .	146
7.2	Setup der xAffect-Software für das Auction Game . . . . .	147
7.3	Ablauf einer einzelnen Session . . . . .	150
7.4	Zusammenhang von Performance und Spielschwierigkeit für beide Gruppen . . . . .	151
7.5	Durchschnittliche Spielperformance für beide Gruppen im Verlauf der Wochen . . . . .	157
7.6	Anwendung von Emotionsregulation in Versuchs- und Kontrollgruppe vor und nach der Studie . . . . .	159

7.7	Reaktionszeiten für beide Gruppen für neutrale und negative Wörter . . . . .	162
8.1	Prinzip des entwickelten Arbeitsgedächtnistrainings für den Fall $n = 3$ . . . . .	166
8.2	Setup der xAffect-Software für das adaptive Arbeitsgedächtnistraining . . . . .	168
8.3	Versionen der $n$ -Back-Task zur Evaluation der Feedbackanzeige . . . . .	169
8.4	Mittlere Trainingsperformance pro Runde nach Quartilsrängen	173
8.5	Relative Herzratenveränderung in Abhängigkeit von der Trainingsperformance . . . . .	174
8.6	Algorithmus zur Adaption der Trainingsschwierigkeit . . .	176
8.7	Methoden zur der Trainingsergebnisse . . . . .	178
8.8	Ablauf der Systemevaluation . . . . .	179
8.9	Adaption zu Beginn des Spiels . . . . .	180
8.10	Verlauf von Performance und Schwierigkeitsgrad über die Sessions . . . . .	181
8.11	Ergebnisse des GEQ-Fragebogens . . . . .	183
8.12	Performanceveränderung für beide Gruppen von Kalibrierungssession zu Abschlussession . . . . .	185

# Tabellenverzeichnis

2.1	Frequenzbänder für die HRV-Analyse . . . . .	34
2.2	Kontingenzmatrix für Klassifikationsergebnisse . . . . .	44
4.1	Einordnung der ausgewählten Messsysteme . . . . .	78
5.1	Auszahlungsbetrag nach erreichtem Rang . . . . .	95
5.2	Kardiovaskuläre Features für die Datenanalyse . . . . .	99
5.3	EDA-Features für die Datenanalyse . . . . .	100
5.4	Maus-Features für die Datenanalyse . . . . .	101
5.5	Mittelwerte und Standardabweichung für Valenz- und Erre- gungsratings vor und nach dem Spiel . . . . .	102
5.6	Minimale Zeitfenster für die Analyse von HRV-Frequenz- bändern in Abhängigkeit von der Untergrenze des Frequenz- bands . . . . .	104
5.7	Ergebnisse der ANOVA für die jeweiligen experimentellen Bedingungen . . . . .	105
5.8	Fisher's discriminant Ratio für alle Features für die unter- schiedlichen Fensterlängen . . . . .	106
5.9	Mittlere prozentuale Veränderung der EDA-Features wäh- rend des Spiels im Vergleich zur Baseline (Standardabwe- ichung in Klammern) . . . . .	112
5.10	Klassifikationsgenauigkeiten der verschiedenen Klassifikato- ren für einzelne physiologische Signale . . . . .	119
5.11	Klassifikationsgenauigkeiten der verschiedenen Klassifikato- ren für unterschiedliche Signalkombinationen . . . . .	119
6.1	Anteil der akzeptierten Daten sowie das dazugehörige F- Maß für unterschiedliche Schwellwerte für Aktivität und Herzrate . . . . .	139



7.1	Mittelwerte und Standardabweichung für die abhängigen Variablen . . . . .	153
7.2	Univariate Tests für die abhängigen Variablen . . . . .	154
7.3	Regression über die Abhängigkeit der Performance von der Herzrate der Versuchspersonen . . . . .	156
8.1	Konfiguration der Trainingsschwierigkeit im Rahmen der Kalibrierungsstudie . . . . .	172
A.1	Mittelwerte und Standardabweichung der HRV-Parameter für unterschiedliche Fenstergrößen über alle Versuchspersonen	223

# Abkürzungsverzeichnis

ANS	Autonomes Nervensystem
DMS	Dehnmessstreifen
EDA	Elektrodermale Aktivität
EEG	Elektroenzephalogramm
EKG	Elektrokardiogramm
ERQ	Emotion Regulation Questionnaire
FDR	Fisher's Discriminant Ratio
FN	False Negative
FP	False Positive
FSR	Force Sensing Resistor
GEQ	Game Experience Questionnaire
GUI	Graphical User Interface
HF	High Frequency
HR	Herzrate
HRV	Herzratenvariabilität
IAPS	International Affective Picture System
kNN	k-Nearest-Neighbour
LDA	Lineare Diskriminanzanalyse
LF	Low Frequency
NS-SCR	Non-Specific SCR
NV	Negativier Vorhersagewert
OSEA	Open Source ECG Analysis Algorithm
PANAS	Positive Affect Negative Affect Scale

PAT	Pulse Arrival Time
PEP	Pre-Ejection Period
PNS	Peripheres Nervensystem
PPG	Photoplethysmographie
PR	Präzision (positive Prädiktivität)
PTT	Puls-Transit-Zeit
QDA	Quadratische Diskriminanzanalyse
RBF	Radial Basis Function
RE	Recall (Sensitivität)
SAM	Self-Assessment Manikin
SCL	Skin Conductance Level
SCR	Skin Conductance Response
SD	Standardabweichung (Standard deviation)
SFFS	Sequential Floating Forward Selection
SFS	Sequential Forward Selection
SP	Spezifität
SUS	System Usability Scale
SVM	Support-Vector-Machine
SWT	Standard Widget Toolkit
TN	True Negative
TP	True Positive
UAR	Unweighted Average Recall
ULF	Ultra Low Frequency
VLF	Very Low Frequency
WAR	Weighted Average Recall
ZNS	Zentrales Nervensystem

# A Werte der HRV-Parameter

	15 s	30 s	300 s
SDNN <sub>LA</sub>	39,840 (20,152)	47,041 (23,506)	51,282 (17,577)
SDNN <sub>HA</sub>	26,805 (16,356)	30,455 (15,765)	40,218 (16,962)
RMSSD <sub>LA</sub>	34,403 (20,022)	36,521 (19,877)	35,824 (17,358)
RMSSD <sub>HA</sub>	21,900 (16,143)	23,034 (15,870)	24,493 (15,514)
pNN12 <sub>LA</sub>	0,619 (0,221)	0,627 (0,205)	0,623 (0,187)
pNN12 <sub>HA</sub>	0,405 (0,258)	0,417 (0,280)	0,427 (0,258)
pNN20 <sub>LA</sub>	0,476 (0,238)	0,477 (0,224)	0,463 (0,209)
pNN20 <sub>HA</sub>	0,264 (0,231)	0,271 (0,245)	0,281 (0,231)
pNN50 <sub>LA</sub>	0,147 (0,158)	0,159 (0,162)	0,156 (0,140)
pNN50 <sub>HA</sub>	0,078 (0,108)	0,077 (0,108)	0,077 (0,108)
SD1 <sub>LA</sub>	25,942 (13,970)	27,194 (14,140)	25,429 (12,313)
SD1 <sub>HA</sub>	16,644 (11,980)	17,023 (11,243)	17,430 (10,956)
SD2 <sub>LA</sub>	47,517 (25,146)	58,937 (31,186)	67,444 (22,704)
SD2 <sub>HA</sub>	32,380 (20,117)	38,506 (19,912)	53,606 (22,470)
SD1/SD2 <sub>LA</sub>	0,578 (0,218)	0,493 (0,197)	0,370 (0,117)
SD1/SD2 <sub>HA</sub>	0,550 (0,301)	0,456 (0,224)	0,321 (0,116)
LF <sub>LA</sub>	-	0,321 (0,141)	0,274 (0,103)
LF <sub>HA</sub>	-	0,373 (0,170)	0,234 (0,115)
HF <sub>LA</sub>	-	0,186 (0,104)	0,119 (0,065)
HF <sub>HA</sub>	-	0,190 (0,097)	0,115 (0,073)
LF/HF <sub>LA</sub>	-	3,030 (3,209)	3,083 (2,159)
LF/HF <sub>HA</sub>	-	3,077 (3,974)	3,208 (2,893)

**Tabelle A.1:** Mittelwerte und Standardabweichung der HRV-Parameter für unterschiedliche Fenstergrößen über alle Versuchspersonen



Diese Arbeit untersucht, wie herkömmliche Lernanwendungen um Informationen über den emotionalen Erregungszustand eines Nutzers erweitert werden können. Den Benutzer zu jedem Zeitpunkt des Lernens auf einem möglichst optimalen Erregungsniveau zu halten, wirkt sich positiv auf den Lernerfolg und im Zuge dessen auch auf die Motivation des Lernenden aus. Da während des Lernens sowohl auf Nutzer- als auch auf Systemseite eine Anpassung erfolgen kann, werden beide Aspekte in dieser Arbeit beleuchtet.