Investigating the Prerequisites for a robust Neurotutor:
The Detection of mixed User States containing Working Memory Load, Affective Valence and Affective Dominance

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Sebastian Grissmann, Mag. rer. nat.

aus Salzburg (Österreich)

Tübingen

2016

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Prüfung: 10.02.2017

Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Peter Gerjets
2. Berichterstatter:	Prof. Dr. Augustin Kelava

for

The Sun

Thank you for granting us continual existence

through your inconceivable power

and not eating us yet.

Acknowledgments

Summary

Intelligent tutoring systems are software environments that aim to simulate a human tutor. While current systems show effectiveness comparable to human tutors, they still suffer from the 'assistance dilemma'. This drawback refers to the inability to infer the ongoing user state which can lead to situations where the system provides no or inadequate support. To alleviate this situation, user state detection has been implemented in some systems. However, at the current time, only behavioral indicators are used to infer the ongoing user state. Such overt behaviors are not specific enough to provide a detailed representation of the user state. This is the reason why I suggest to investigate the potential use of the electroencephalogram to infer the ongoing user state. This combination of an intelligent tutoring system and an EEG-based user state detection is called a neurotutor. EEG-based user state detection usually focusses on narrow user states which can be detected in controlled lab environments. I assume that real-life environments like a classroom evoke complex user states which consist of multiple different components. I therefore propose a three component framework that enables the tracking of different processes that are active during a complex user state. The first two studies focus on the separation of working memory load and affective valence in a highly controlled setting with the use of established measures from classical neuroscience. I found that measures used to infer working memory load can be used to track changes in working memory load under different affective valence. Furthermore, I found that said measures were also sensitive to changes in affective valence. Surprisingly, I found that measures used to infer affective valence were not sensitive to changes in affective valence under working memory load. Additional analyses revealed that working memory load and affective valence can be automatically detected with accuracies sufficient for the use in a neurotutor. The third study successfully replicated the findings from the first two studies in a more realistic, although less controlled setting. A simplified learning game was used to induce the complex user state of perceived loss of control that simultaneously evoked cognitive as well as affective processes. With the help of the framework I was able to integrate the findings from three different studies that all analyzed the same dataset. This would not have been possible without an adequate theoretical framework.

Zusammenfassung

Intelligente Tutorensysteme sind EDV-Programme, welche versuchen einen menschlichen Tutor zu simulieren. Obwohl derzeit erhältliche Systeme ähnlich effektiv sind wie menschliche Tutoren, leiden sie immer noch unter dem ‚Assistenzdilemma'. Dies referenziert auf die Unfähigkeit den aktuellen Nutzerzustand zu erkennen, was dazu führen kann, dass das System keine oder inadäquate Unterstützung anbietet. Um diesen Umstand zu beseitigen wurde in manchen Systemen bereits eine Nutzerzustandserkennung implementiert. Derzeit werden jedoch nur beobachtbare Verhaltensweisen verwendet, um den aktuellen Nutzerzustand abzuleiten. Solch offenkundige Verhaltensweisen sind jedoch nicht spezifisch genug, um ein detailliertes Bild des Nutzers zu erfassen. Deshalb schlage ich vor, dass man das Elektroenzephalogramm verwenden soll, um den aktuellen Nutzerzustand zu erkennen. Diese Kombination eines intelligenten Tutorensystems mit einer EEG-basierten Zustandserkennung nennt man einen Neurotutor. Normalerweise fokussiert sich EEG-basierte Zustandserkennung jedoch nur auf begrenzte Zustände, welche innerhalb kontrollierter Laborumgebungen erkannt werden. Ich nehme an, dass realistische Settings wie ein Klassenzimmer komplexe Zustände hervorrufen, welche aus mehreren verschiedenen Komponenten bestehen. Daher stelle ich ein Bezugssystem mit drei Komponenten auf, welcher die Zuordnung der verschiedenen Prozesse, die während eines komplexen Nutzerzustandes aktiv sind, ermöglicht. Die ersten beiden Studien befassen sich mit der Separation von Arbeitsgedächtnisbelastungen und affektiver Valenz durch die Verwendung von etablierten EEG Maßen in einem hoch kontrollierten Setting. Dabei habe ich herausgefunden, dass Maße, welche zum Erkennen von Arbeitsgedächtnisbelastungen verwendet werden, auch geeignet sind, um Veränderungen in der Belastung des Arbeitsgedächtnisses unter gleichzeitiger emotioneller Stimulation zu erkennen. Außerdem habe ich entdeckt, dass die erwähnten Maße auch sensitiv gegenüber Veränderungen in der affektiven Valenz sind. Überraschender Weise hat sich herausgestellt, dass Maße, welche zum Erkennen von affektiver Valenz verwendet werden, nicht sensitiv gegenüber Veränderungen in der affektiven Valenz sind, wenn gleichzeitig das Arbeitsgedächtnis belastet wird. Zusätzliche Analysen haben aufgedeckt, dass Arbeitsgedächtnisbelastungen und affektive Valenz mit einer Genauigkeit erkannt werden kann, welche für den Gebrauch in einem Neurotutor ausreichend sind. In der dritten Studie konnten die Ergebnisse der ersten beiden Studien erfolgreich in einem relevanteren, aber weniger kontrollierten Kontext repliziert werden. Dabei wurde ein vereinfachtes Lernspiel verwendet, um einen wahrgenommenen Kontrollverlust zu induzieren. Durch Verwendung des Bezugsrahmens war es mir möglich die Ergebnisse von drei verschiedenen Studien zu integrieren, welche alle denselben Datensatz analysierten. Dies wäre nicht ohne adäquaten theoretischen Rahmen möglich gewesen.

Contents

Abbreviations (in alphabetical order)

BA   Brodmann Area

BCI   Brain-Computer Interface

CAI   Computer-Assisted Instruction

CLT   Cognitive Load Theory

EDA   Electrodermal Activity

EEG   Electroencephalogram

EF   Executive Function

ERD/S   Event-Related DeSynchronization

ERP   Event-Related Potential

FAA   Frontal Alpha Asymmetry

fMRI   functional Magnetic Resonance Imaging

fNIRS   functional Near-Infrared Spectroscopy

IADS   International Affective Digitized Sounds

IAPS   International Affective Picture System

IC   Independent Components

ICA   Independent Component Analysis

ITS   Intelligent Tutoring Systems

MCM   Multi-Component Model

MEG   Magnetoencephalography

MOOC   Massive Open Online Courses

LPP   Late Positive Potential

PAD   Pleasure-Arousal-Dominance Model

RQ   Research Question

SAM   Self-Assessment Manikin

SMA   Supplementary Motor Area

SSVEP   Steady-State-Visually-Evoked Potential

WM   Working Memory

WML   Working Memory Load

Overview

In this doctoral thesis I investigated the potential of the electroencephalogram (EEG) for the detection of complex user states that could arise during realistic human-machine interaction in educational settings.

The first four sections describe the background and motivation of this thesis. In the first section I will give an overview on Intelligent Tutoring Systems (ITS), their historical development as well the effectiveness and real-life applicability of ITS. It will describe different tutoring designs and a prominent example of a current ITS, yet it will also highlight the main limitation of current ITS. This will lead toward the second section of this thesis which will present a short introduction into the novel research field of neuroeducation. There I will outline the application of the EEG for user state detection, an approach that could have the potential to address an important issue in the development of ITS, the so-called assistance dilemma. The section will further contain a short overview of the EEG, the main method used in this thesis. The third section will describe a newly adopted approach in the field of neuroeducation, which focuses on the automatic detection of user states and is called a brain-computer interface (BCI). This will bring readers to the concept of neurotutors, describing recent efforts toward the development of such systems, but also emphasizing current challenges for the development of robust systems which can be used in natural environments like classrooms. As the classroom represents an environment that could elicit complex user states, containing different mixtures of cognitive processes as well as affective processes, the fourth section will focus on the description of such complex user states. In this section of the thesis the main components of a proposed framework will be described, including related theoretical frameworks, paradigms used to investigate such complex user states, induction methods as well as EEG measures used to infer specific components of the proposed framework.

The next three sections will focus on the research conducted as part of this thesis. The fifth section of this thesis will center on the aim of this research and the corresponding research questions (RQs). After that, the sixth section will contain three manuscripts, which represent the main body of empirical evidence of this doctoral thesis. This section will also contain additional sub-sections which will present results from the analysis of subjective measures used in the first study, the automatic detection

of affective valence in a complex user state, additional analyses related to the third study as well as results from a related study that analyzed the same dataset that was used in the third study. All this additional content has not been included in the submitted manuscripts. Following this, the seventh section is an attempt to integrate and discuss all relevant findings.

The last two sections will present an outlook for future research. In the eight section, the next necessary steps, which are required to bring neurotutors into natural environments, will be discussed. Finally, the ninth section will focus on the practical relevance, briefly describing the enormous potential for neurotutors in the educational context and for lifelong learning.

1. Intelligent Tutoring Systems

ITS are computer systems designed to simulate a human tutor with the aim to provide access to high quality education to each and every student, and therefore likely to reform the entire education system sometime in the future. While ITS are expensive to develop (Murray, 1999), they could help create engaging learning situations using customized instruction and feedback, usually via simulation of discourse patterns and pedagogical strategies of a human tutor (Russell, Moran, & Jordan, 1988).

1.1 A short History of Intelligent Tutoring Systems

First attempts to utilize machines for instructional purposes date back to 1924. The Pressey machine was a metal box with several buttons and a window that was used to present questions. Users could select their response by pressing a button and the machine then provided immediate feedback concerning their performance by counting and then displaying the score (Shute & Psotka, 1985).

Based on advances in computer science in the 1960s and 1970s many new computer-assisted instruction (CAI) programs were initiated in countries like the US, the UK, as well as Canada (Chambers & Sprecher, 1983). While CAI was gaining interest, Carbonell (1970) suggested that computers could act more like a teacher rather than a mere tool. Based on this new perspective on the potential of educational machines, intelligent computer-assisted instruction, currently known as ITS would emerge. This new approach focused on the use of computers to intelligently support students in their learning progress. While CAI programs were based on a behaviorist perspective, grounded in learning based on Skinner's theories (Ormrod, 2011), ITS are based on interdisciplinary developments in cognitive psychology, computer science, and especially artificial intelligence (Larkin & Chabay, 1992).

Since computers were still expensive and not as available as expected at the end of the 1970s, interest in ITS technologies was fading. However, the microcomputer revolution in the early 1980s, changed this situation by reducing costs and thereby helped to revive development of ITS (Anderson, 1986). This can be seen as the foundation of current ITS.

1.2 Effectiveness and real-life Applicability of Intelligent Tutoring Systems

Effectiveness of ITS is commonly assessed by comparing them with human tutors. While different ITS vary in their effectiveness, a recent review by VanLehn (2011) found similar effect sizes when comparing ITS with human tutors in general. While most studies using ITS are still conducted in laboratory settings, some studies already highlight the potential of ITS in real-life settings like classrooms.

For instance, a study by Koedinger, Anderson, Hadley, & Mark (1997) investigated the applicability of ITS in a large-scale approach, involving multiple urban high schools. The authors found that ITS could help to improve performance on standardized math tests by as much as fifteen percent.

1.3 Different Tutoring Designs

Current ITS track the progress of students by comparing their problem solving approach with those of experts. There exist two different approaches to do this. Firstly, there are approaches like the Cognitive Tutor that implement pure knowledge tracing (Anderson & Corbett, 1995). Such systems track the progress of a student using an expert knowledge module and a student model module (Nwana, 1990). The expert knowledge module contains all the necessary steps toward a problems solutions. The student model module can be seen as an overlap of the expert knowledge module. Potential deviations from the expert knowledge module are used to indicate errors. While such systems are able to identify when a student has troubles with a specific problem, they are not capable to identify potential causes for those troubles.

This led to the design of dialogue-based affective tutors. Such systems try to detect the current learner state with the intent to identify moments where conceptual change is possible (e.g., during states of confusion; D'Mello, Lehman, Pekrun, & Graesser, 2014). Conceptual change refers to the change of concepts when one is getting rid of misconceptions (Duit & Treagust, 2003). One prominent example for such an affective tutoring system is the AutoTutor (Graesser et al., 2004).

1.4 The AutoTutor: An Example of an Affective Intelligent Tutoring System

Currently available ITS have been designed to help students learn amongst other topics, geography, medical diagnosis, computer programming, mathematics and physics. One well known example is the AutoTutor (Graesser, 1999) which was designed to help college students in an introductory computer literacy course. In contrast to the Cognitive Tutor, The AutoTutor is focusing on natural language dialogue. It is formulating dialogue responses using production rules based on learner contributions as well as the detection of the current learner state (Graesser et al., 2004).

The current version is able to detect different cognitive as well as affective states via conversational cues, gross body language and facial features (D'Mello et al., 2008; D'Mello, Picard, & Graesser, 2007). Furthermore, the AutoTutor can provide emotional feedback via an embodied emotional agent that uses facial expressions (e.g., showing delightful facial expressions, when a student solves a problem) as well as modulated speech (e.g., changing intonation based on the quality of the current answer from a student). **Figure 1** shows a sample screenshot of the AutoTutor interface and **Figure 2** shows the different sensors used for detection of the current learner state used in the student model.
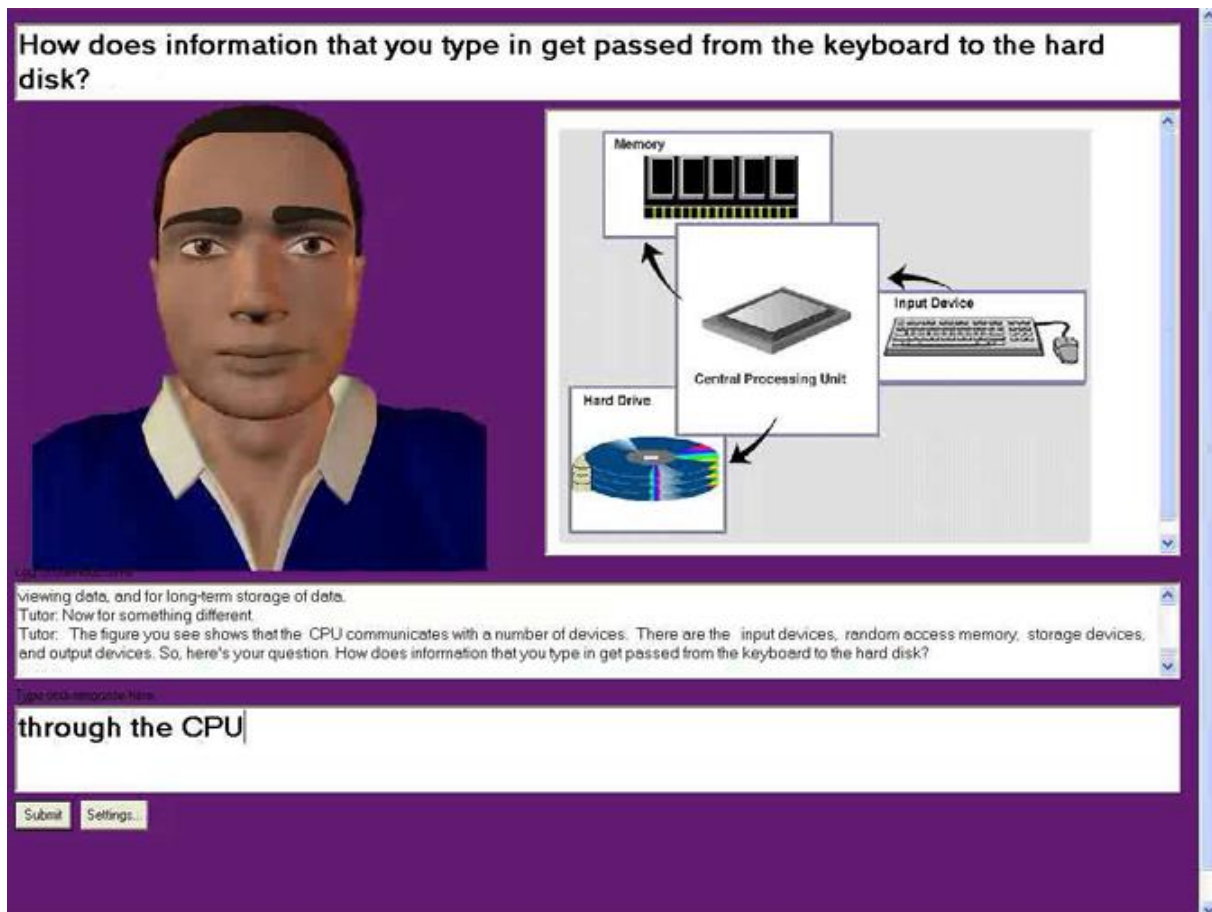
Figure 1. The AutoTutor interface. At the top the current question is displayed. The embodied emotional agent can be seen at the left side and the right side shows an illustration related to the current question. In the middle of the screenshot one can see the last responses from the AutoTutor. The input window at the bottom of the screen can be used to enter responses (https://commons.wikimedia.org/wiki/File:AutoTutor.png).
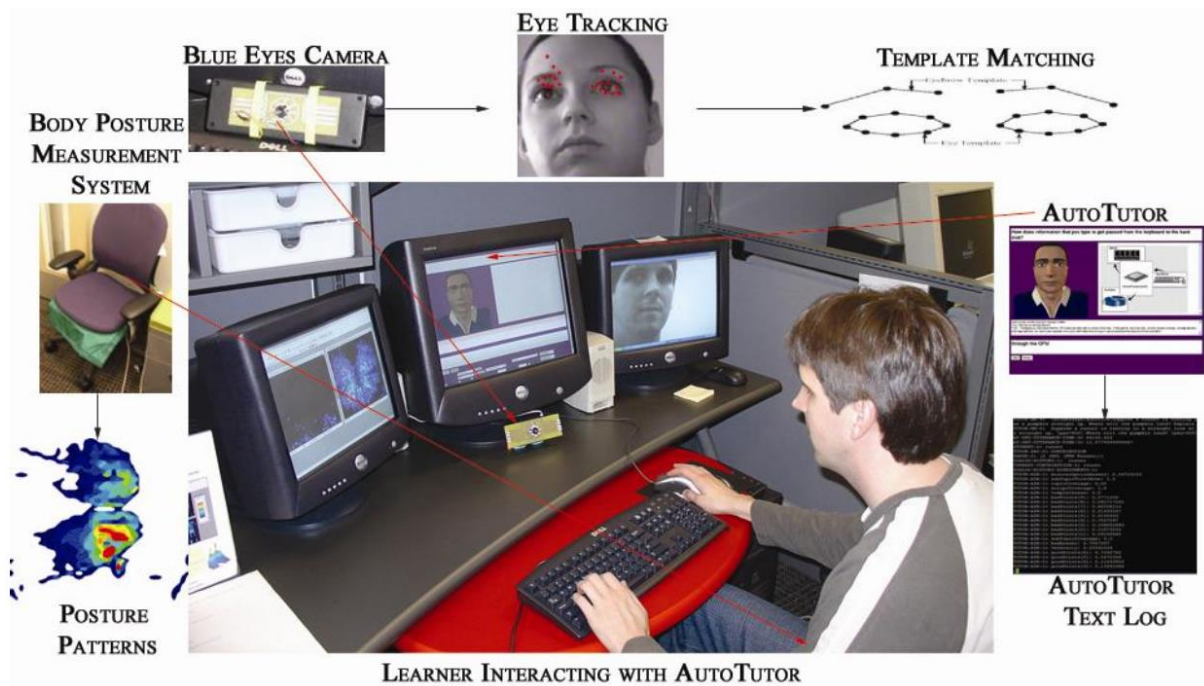
Figure 2. Sensors used for detection of learner state in the AutoTutor. The AutoTutor uses pressure pads to detect the current posture of the user as well as eye-tracking to extract facial features. Furthermore, the AutoTutor analyzes the responses of the user via a text log (D'Mello et al., 2008).

While the detection of different user states is definitely an improvement over the black box approach implemented in approaches that only use knowledge tracing, such systems still cannot infer why a user is experiencing a certain state. If one looks at the example of a frustrated student. It might be that he or she is frustrated because the task is too difficult and the student experiences mental overload. However, it might also be the case that the student is bored because the task is too easy. Both cases would require opposing interventions to get the student back on the optimal learning track. Furthermore, currently implemented sensors like pressure pads and cameras are not specific enough to make reliable judgments about the ongoing user state. One can imagine a case where a student furrows his or her brows. This could indicate a concentrated state, frustration or even anger. The ambiguity of such sensor measurements leads us to the core of the 'assistance dilemma', which is the question how learning environments should balance assistance giving and withholding to achieve optimal student learning (Koedinger & Aleven, 2007). On one hand, if the system provides help and the student is concentrated, then this would interfere with deep learning. On the other hand, if the system does not provide help and the student

is frustrated, then this could lead to task disengagement. These examples shows that a meaningful modeling of the current user state needs to account for cognitive as well as affective aspects, something that cannot be done with currently implemented sensor modalities. Therefore I propose that we need to find sensor modalities that could allow such improved user state detection, something that might be possible with methods from the novel research field of neuroeducation.

2. The novel Research Field of Neuroeducation

The investigation of brain activity in the context of education and learning has received increased interest in the last couple of years. This is reflected in the formation of the new research field called neuroeducation, also known as educational neuroscience (Ansari, Smedt, & Grabner, 2011; Sigman, Peña, Goldin, & Ribeiro, 2014). Neuroeducation aims to apply neuroscientific methods to study the neural causes and correlates of cognitive as well as affective processes, which are relevant to knowledge acquisition. This interdisciplinary synergy of fields like neuroscience, educational science as well as cognitive psychology aims to produce novel findings that help to better understand and thereby potentially improve educational practice (Ansari & Coch, 2006).

To give a concrete example, Kim, Lee, Chung, & Bong (2010) found that for low-competence learners norm-referenced feedback (i.e. comparing a student with other students in the class) can activate brain areas associated with negative emotions, even if the feedback was not negative. This finding, only possible through the use of neuroscientific methods, indicates that normative assessment should be used with caution, especially for low-competence students, since this feedback method might evoke negative emotions regardless of the actual performance.

Most studies in neuroeducation focus on basic processes, which has led to some criticism toward the field (Bowers, 2016). However, recently there have been more intervention-like approaches that could help to create individualized educational environments that support learners coping with the increasing demands of our educational systems (Walter, Schmidt, Rosenstiel, Bogdan, & Gerjets, 2013). Such approaches typically use the EEG to infer the ongoing user state due to its properties which will be described in the next section.

2.1 An Overview of the EEG

Biosignals like the EEG could offer some advantages for the detection of covert user states when compared to movement analysis and analysis of facial expressions (Broek et al., 2010). For instance, movement analysis is rather crude like in the case of pressure pads or requires cost-intensive motion tracking sensors and ample free space. This might not be very practical for target scenarios like classrooms. The

analysis of facial expressions can be difficult, especially in dark environments or if the user is changing position. Furthermore, most current approaches for the detection of facial expressions are based on exaggerated expressions staged by actors and therefore it seems difficult to determine performance of such approaches under realistic conditions (Zeng, Pantic, Roisman, & Huang, 2009). In contrast to that, the EEG is difficult to manipulate intentionally and could therefore provide objective measures for the detection of the current user state (Broek et al., 2010). Accordingly, the EEG could help to solve the assistance dilemma by revealing aspects of the ongoing user state that are not accessible by analysis of overt behaviors alone. This could help to create a more differentiated picture if the current learner state.

The following section will provide a brief summary of key elements in EEG-based research (for a more detailed description see Niedermeyer, Schomer, & da Silva, 2011). The EEG is a relatively cost-efficient brain imaging method when compared to other methods like functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG) or functional near-infrared spectroscopy (fNIRS). Recordings of the EEG are derived from electrodes that are placed on the scalp and measure voltage fluctuations due to ionic currents resulting from synchronous oscillations of large numbers (>10.000) of neurons in the brain. The activity of neural assemblies can change extremely fast and therefore the EEG can record with high temporal resolution, which is in the range of milliseconds. The EEG provides a direct measure of neural activity and not indirect, metabolic measures like the fMRI and fNIRS. It has been thoroughly studied since first recordings have been conducted by Berger (1929). Due to this vast knowledge base, which contains investigations of various different mental states and processes, it could enable the unobtrusive real-time monitoring of multiple different affective as well as cognitive states and processes which are reflected in changes of oscillatory activity. Furthermore, cost-efficient EEG systems are already available, which is very important in the context of ITS, since large scale applications in a school context require that the necessary hardware is affordable (see **section 8.5**). However, there are also some drawbacks. The EEG mainly records activity from pyramidal cells with a specific orientation toward the scalp and therefore activity from subcortical regions can be hard to detect using this method. Furthermore, the EEG has a rather bad spatial resolution, which is in the range of cm². Additionally, the EEG is susceptible to artifacts, which can be induced by movements, muscle

activity or unrelated electrical signals. While systems with dry electrodes that do not require extensive preparation are already available (Yeung et al., 2015), traditional EEG systems still take considerable time before they can be used (mounting of the cap and positioning of the electrodes). The analysis of EEG signals is commonly divided into two main categories, approaches based on the time-domain and approaches based on the frequency-domain.

2.1.1 Time-Domain based Approaches for the EEG

EEG approaches that focus on the time-domain investigate the impact of different mental states and processes on so-called event-related potentials (ERPs; for an detailed description see Kappenman & Luck, 2012). ERPs are deflections in the recorded EEG signals which are time- and phase-locked to a specific stimulus event (e.g., presentation of a picture stimulus). EEG recordings also contain noise due to artifacts not related to the event of interest (e.g., background noise in the EEG). To reduce these unwanted aspects of the signal, multiple repetitions of the same stimulus are commonly used to get an averaged, also called prototypical response.

The averaged ERP response is called the ERP curve, which shows positive and negative deflections with regard to a baseline. Specific components of the ERP curve are categorized by a letter in front of the description ('P' for positive deflections and 'N' for negative deflections). Furthermore, the approximate latency after stimulus onset is indicated by a number (e.g., the P300 is a positive deflection occurring on average 300 ms after stimulus onset). ERP curves are further grouped into early and late potentials. Early potentials are loosely defined as occurring before 300 ms and related to automatic processes in response to a specific stimulus (e.g., automatic perceptual processing of stimulus properties). ERPs occurring later than 300 ms after stimulus onset are called late potentials and correspond to higher level processes in the brain. In realistic settings averaging across several instances of the same stimulus is rarely feasible, since it would require to present the same stimuli repeatedly. Although single-trial ERP classification approaches have been shown to provide comparable classification accuracies in certain cases (Blankertz, Lemm, Treder, Haufe, & Müller, 2011) and recently developed approaches even work without being time-locked to a stimulus (Krumpe, Walter, Rosenstiel, & Spüler, 2016), ERP-based approaches cannot

provide continuous, fine-graded measurements and therefore the usefulness of ERPs for adaptive learning environments is likely to be limited.

2.1.2 Frequency-Domain based Approaches for the EEG

The frequency-domain might offer an alternative to time-domain based approaches that can provide continuous measurements without the need for repetitive external stimulation. Analyses in the frequency domain use two, fundamentally different power extraction methods, resulting in evoked or induced responses (Tallon-Baudry & Bertrand, 1999). An evoked frequency response applies a frequency transform (e.g., via fast Fourier transform) on the averaged EEG signal and thereby creates a frequency-domain representation of the ERP signal. In an induced frequency response the frequency transform is applied on single trials of EEG data before averaging and therefore also captures parts of the EEG signal which are not phase-locked to the stimulus.

Frequency responses extracted from the EEG signal are grouped into distinct frequency bands which are assumed to be roughly related to different types of mental processes and states (Krause, 2003). Oscillations within the lowest frequency range are part of the delta band which ranges from 0.5 to 4 Hz. Delta band activity is commonly associated with slow wave sleep, but has also been found in babies during wake phases (Torres & Anderson, 1985). The second slowest frequency range is called the theta frequency band, ranging from 4 to 8 Hz. Activity in the theta band is related to a number of cognitive processes and is particularly relevant in the context of working memory load (WML; Klimesch, 1996; see also **section 4.3.3.2.1**). The alpha frequency band ranges from 8 to 13 Hz, displaying peak activity around 10 Hz over parietal-occipital brain areas. It was one of the first frequency bands identified in the human brain and therefore it is also the most thoroughly investigated frequency band (Berger, 1929). Activity in the alpha band is related to cortical inactivity, induced, for example, due to a closing of the eyes. In contrast, opening of the eyes reduces alpha band activity and is therefore called alpha-blocking. Alpha activity is related to a wide range of different mental processes. In central areas, alpha-blocking has been shown to be related to sensorimotor activity (Pfurtscheller & Brunner, 2006). Since oscillatory activity over central areas relates to different mental processes than in other brain

regions it is referred to as rolandic mu-rhythm rather than alpha rhythm. Frontal alpha asymmetries are related to different affective states (Allen, Coan, & Nazarian, 2004) and parietal alpha activity is related to changes in WML (Klimesch, 1999). All of these different phenomena will be discussed later in greater detail (see **section 4.5** for motor related functions, **section 4.4.3.2** for affect related functions and **section 4.3.3.2.2** for workload related functions). The beta frequency band ranges from 13 to 30 Hz and is likewise related to multiple different processes. For instance, at central areas beta activity is associated with somatosensory responses. Finally, the gamma frequency band represents the highest frequencies recordable via the EEG, comprising oscillations between 30 and 200 Hz. These types of high-frequency oscillations are assumed to be involved in the integration of information in different sensory and non-sensory cortical networks. However, muscle artifacts also tend to exhibit activity in higher frequencies which makes this frequency band especially prone to artifact contamination. Approaches based on the time-domain as well as the frequency-domain usually tend to consider signals only as they are recorded at the scalp or electrode level. Yet, such approaches fail to account for the fact that the electrodynamic properties of brain matter can lead to volume conduction which can spread neural signals across large brain areas. It is therefore assumed that EEG electrodes record a mixed sum of underlying cortical activity, not only originating from brain regions directly under the electrode site. To differentiate between these different sites of origin and to detect cortical sources related to specific brain functions, EEG-based source localization methods have been developed recently.

2.1.3 Source Analysis of the EEG Data

Recently developed source localization approaches may add to classical approaches based on the time-domain as well as the frequency-domain. Such approaches try to identify cortical sources that create activity recorded at the scalp level. This new approach is based on the observation that changes in activity of small cortical sources can have a significant impact on the EEG via volume conduction. A simulation created by Zeynep Akalin Acar at the Swartz Center for Computational Neuroscience (University of California San Diego) attempts to illustrate this phenomenon. **Figure 3** shows a snapshot of the simulation, composed of two small cortical sources and their corresponding EEG scalp projections.
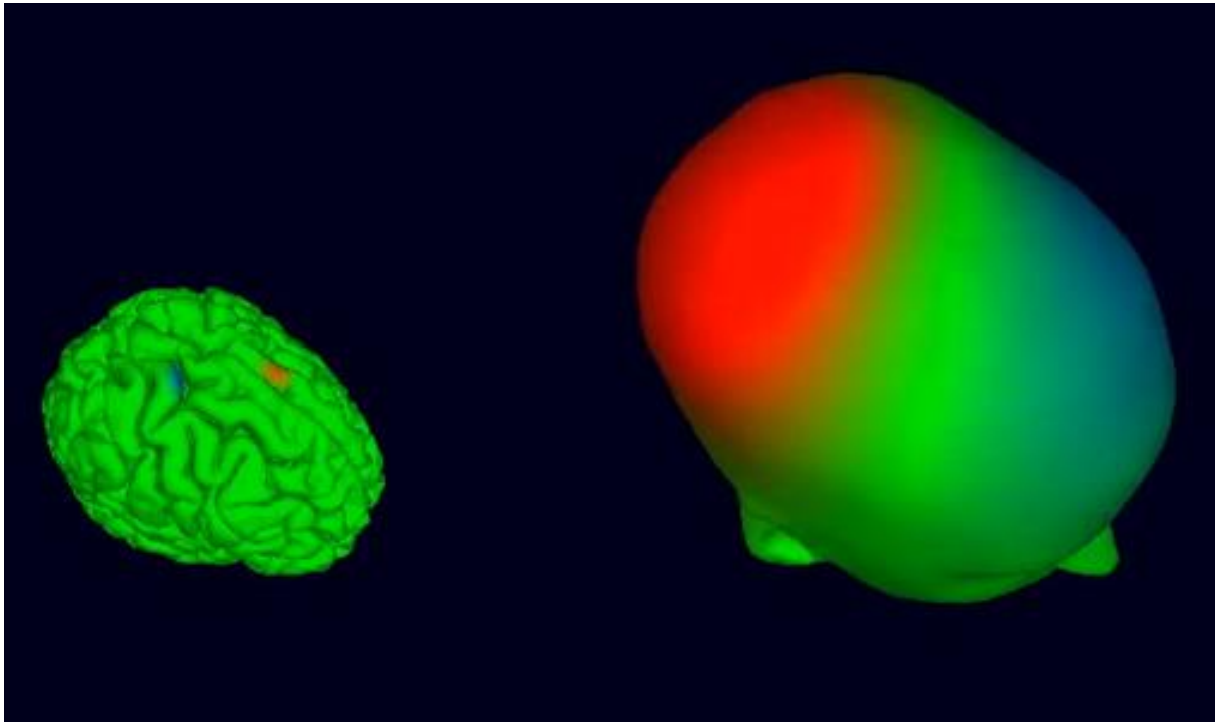
Figure 3. Simulation of two cortical sources and the corresponding EEG scalp projections. The left side shows a red and a blue blob, representing two cm-scale, spatially static cortical EEG sources. These cortical sources are animated with simulated alpha band activity at 9 Hz and 10 Hz respectively. The image on the right displays the summed projections of the two brain sources onto the scalp surface, which is equal to the scalp EEG dynamics produced by these sources (the full video can be accessed via http://www.youtube.com/watch?v=7TP-xrVokXk).

With the help of this simulation it can easily be observed that even small fluctuations in cortical activity can have a large impact on the signals that are recorded via electrodes at the scalp. Since this effect is present in all scalp-based EEG analyses, considering only the signal as it is recorded at the scalp might result in misleading conclusions considering specific cortical activity (Makeig, Bell, Jung, & Sejnowski, 1996).

One approach to get from sensor space to source space is via equivalent dipole modeling based on topographies of independent components (ICs). Independent component analysis (ICA) attempts to separate the scalp signal into multiple statistically independent components, based on the assumption that EEG data is a linear sum of underlying cortical activity. These ICs usually exhibit scalp-projections which are 'dipole-like' (Delorme, Palmer, Onton, Oostenveld, & Makeig, 2012) and

therefore dipole modeling can be used to locate the three-dimensional position of the ICs. This allows an approximate localization of the cortical sources. Multiple different approaches used to estimate these cortical sources exist and are usually selected based on the available computational resources (Oostenveld, Fries, Maris, & Schoffelen, 2011). However, the precision of the location of the cortical sources is in any case limited due to the underdeterminedness of the inverse solution (Onton, Westerfield, Townsend, & Makeig, 2006). This requires that results based on this, rather qualitative method have to be checked for psychological and neurophysiological validity to make them applicable in the field of neuroeducation. While the last section was focusing on different approaches to measures user states with the EEG, the combination with an ITS also requires the automatic use of such measures to detect the ongoing user state.

3. Automatic Detection of User States

The automatic detection of user states via brain imaging techniques like the EEG is the main idea behind BCIs. BCIs were first proposed by Vidal in 1973. They can provide a new information channel directly from the brain which can be used for direct or indirect control of a given technical system (Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002). BCIs are closed-loop systems that work in the following way. They record activity from the brain using brain imaging methods like the EEG. After that features like band power are extracted from these brain signals. The extracted features are then translated into control signals via machine learning algorithms that allow a computer to learn without being explicitly programmed. Such control signals are then used to provide feedback to the user in multiple different user scenarios (see **Figure 4** for an illustration of the main components).
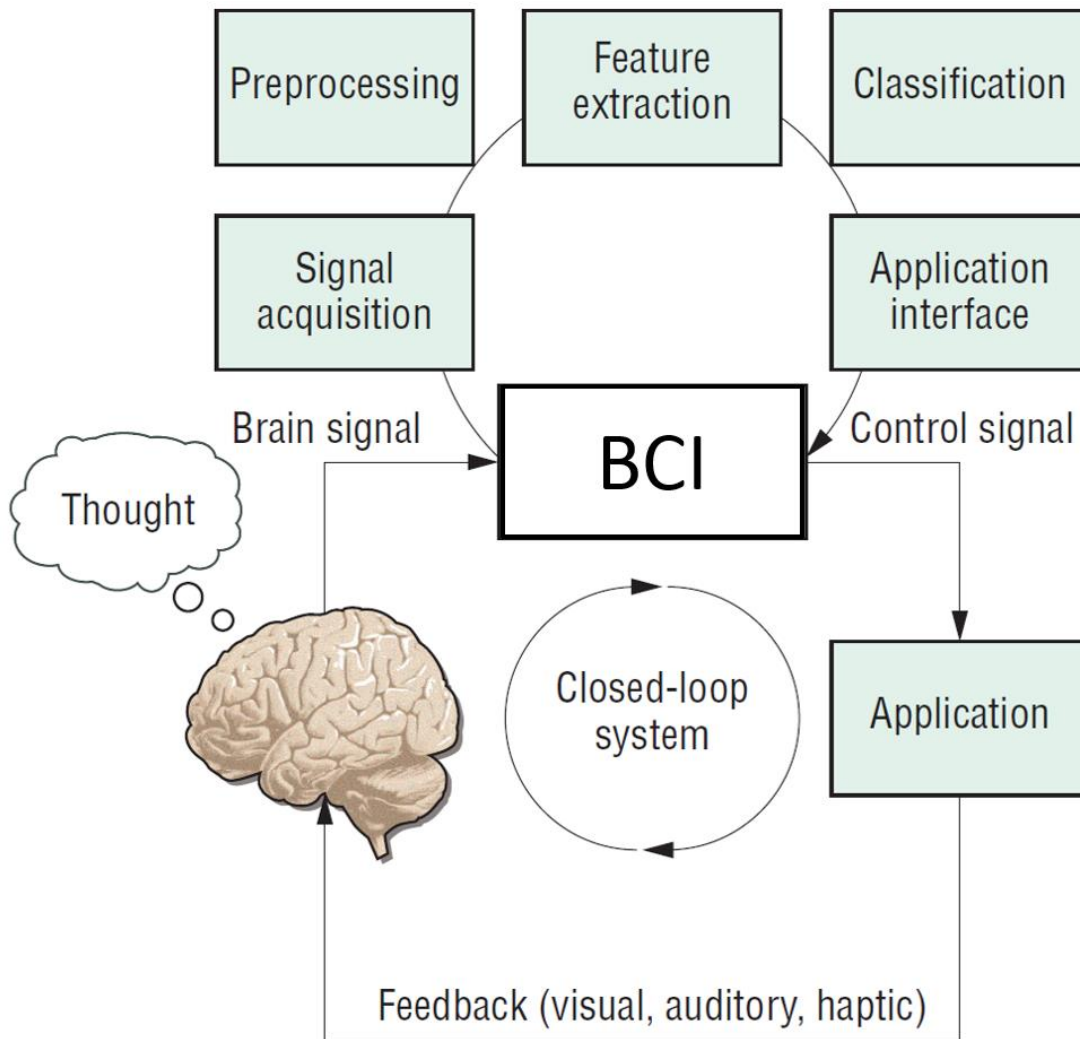
Figure 4. The basic components of a brain-computer interface (BCI) (modified from Nijholt et al., 2008).

Originally, such systems were developed for people that lack voluntary muscle control due to severe impairments in the motor system, resulting from amyotrophic lateral sclerosis or central nervous system stroke. First applications consisted of simple spelling devices that helped to re-establish lost communication, but future BCI systems are designed to be used in a wide range of application scenarios. BCIs have the potential to further replace functions that are lost due to injury or disease (e.g., two-dimensional cursor control; Allison et al., 2011), restore lost functions (e.g., muscle stimulation in paralysis; Daly & Wolpaw, 2008), improve functions (e.g., stroke rehabilitation; Wang et al., 2009) and be used as research tool to study brain functions in general (e.g., optimizing neurocognitive hypothesis testing; Sanchez, Lecaignard,

Otman, Maby, & Mattout, 2016). Finally, they could also help to enhance functions (e.g., measuring attention in distance learning settings; Li et al., 2011).

BCIs can be grouped into three different categories. Firstly, there exist reactive BCIs that require external stimuli to elicit a detectable brain response (e.g., P300 spellers; Krusienski et al., 2006). Secondly, there exist active BCIs which are based on voluntary brain activity like motor imagery (e.g., Brunner, Naeem, & Leeb, 2007). Such systems do not need external stimulation, but in certain cases require extensive training to achieve performance rates which are satisfactory (e.g., BCIs based on slow cortical potentials; Birbaumer, 2006). Finally, there are passive BCIs that combine mental state monitoring with the BCI feedback loop (Zander & Kothe, 2011; Zander, Kothe, Jatzev, & Gaertner, 2010). Passive BCIs utilize the continuous assessment of user states via the EEG in order to provide adaptive feedback. Thus, in contrast to active and reactive BCIs, passive BCIs do not require that users intentionally control the BCI system with their thoughts. Ideally, passive BCIs could improve the flow of information in human-machine systems without any additional mental effort (Zander et al., 2010). Passive BCIs appear especially useful if one wants to assess covert aspects of the current user state, which cannot be inferred based on behavioral cues (Reissland & Zander, 2009). See **Figure 5** for an illustration of a passive BCI that provides adaptive feedback based on an individual learners emotions, which is also called a neurotutor.
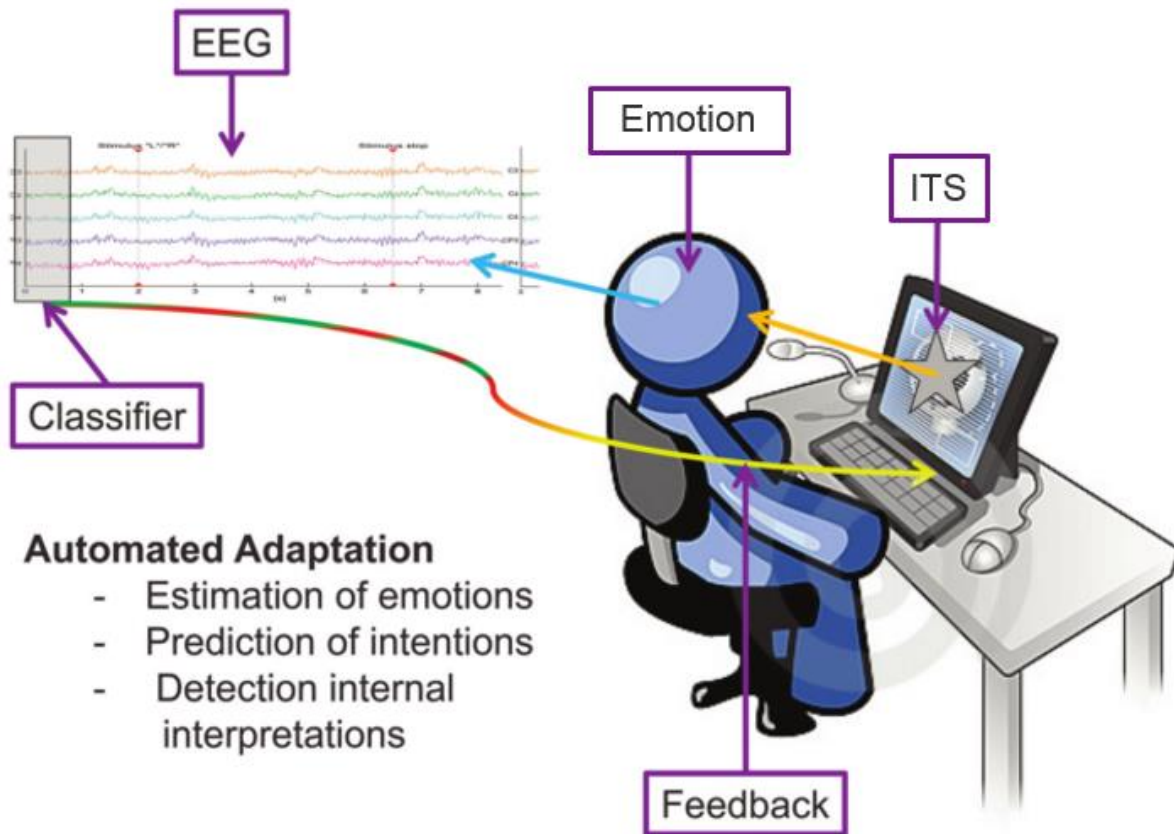
Figure 5. Basic principle behind a neurotutor, a passive BCI system used in an educational context. Users are working on a learning task (e.g., using an intelligent tutoring system). Simultaneously, the brain activity of the users is measured via the electroencephalogram (EEG). A classifier derived from machine learning methods is used to assess their current affective state. This affective state evaluation is fed back into the system to be used for automatic adaption of the learning environment (modified from Zander, 2011).

3.1 Neurotutors

Previous research already hinted at the potential of the use of the EEG in educational settings. For instance, Heraz and Frasson (2007) used brain waves recorded with the EEG to predict the three major dimensions of learner's emotions, pleasure, arousal and dominance (see **section 4.4.1.2** for a more detailed description of these affective dimensions). Emotions were induced via affective pictures and the authors concluded that EEG recordings could be utilized to infer different emotional dimensions in an educational context. A more recent investigation by Wang et al. (2011) used the EEG with the aim to improve massive open online courses (MOOC) by providing additional information about the users while they are studying the course material. In their study they presented video clips and used EEG recordings to detect when students were

confused by the course material. They found weak, but above chance classification rates for EEG-based confusion detection. Nevertheless, their approach was still performing comparable to human observers who monitored body language to infer student's confusion levels. A study by Mostow, Chang, and Nelson (2011) used the EEG in combination with a reading tutor and found that simple and complex sentences can be distinguished with better than chance accuracy. Additionally, they were able to identify which EEG components might be sensitive to which particular lexical features. Azcarraga, Marcos, and Suarez (2014) used EEG for the prediction of academic emotions like frustration, confusion, boredom as well as interest. The authors recorded the EEG of young learners (aged 12 to 16) while answering algebra problems using the Aplusix algebra learning software and found accuracies for emotion prediction as high as seventy-five percent. While these studies show the potential of such an approach, they all suffer from methodological flaws. All of the mentioned studies used cheap consumer equipment to record the EEG, which is known to produce worse results than medical grade equipment, even when recording strong signals like ERPs (Duvinage et al., 2013). Furthermore, all but one study (Azcarraga et al., 2014) used only one or two EEG sensors, which does not allow for appropriate artifact removal. This makes the obtained results highly questionable. Nevertheless these studies can still be seen as a precursors on the way toward a neurotutor, a new and promising application of BCI methodology in the field of neuroeducation (Galway, McCullagh, Lightbody, Brennan, & Trainor, 2015).

A neurotutor combines an ITS with a passive BCI approach and aims to monitor the learner state with the intend to adapt the learning environment based on brain activity. This approach could allow to extract markers for multiple cognitive and affective user states which might be used to adapt and optimize the learning strategy of an individual learner, thereby potentially avoiding the 'assistance dilemma'. When studying a new subject, complex concepts would only be explained when users are experiencing states impeding the learning progress, like frustration or mental overload. Therefore, neurotutors may help to optimize the learning process by constantly directing the learner toward the zone of proximal development (Wass & Golding, 2014).

Some researchers have already shown the potential of such technologies. A study by Szafir and Mutlu (2013) used EEG recordings in a paradigm based on adaptive content review. They monitored student's attention while presenting different

lecture topics taken from an online art lesson and later presented those lessons which received the least attention for a second time. The authors found that adaptively reviewing lesson content improved recall by twenty-nine percent when compared to a baseline measure. Additionally, their approach was able to match recall gains achieved by a full lesson review, but took significantly less time. However, the authors also used a consumer grade headset which does not allow the recording of data with sufficient quality.

A recent study by Spüler et al. (2016) used EEG-based workload detection during addition tasks and found that cognitive workload estimation is possible at an above chance level. Furthermore, they found that their model was able to perform on a between-subjects basis. This is of particular interest, since commonly used approaches require that each single subjects provides training data for calibration of the system. Cross-subject approaches could get rid of this requirement and thereby significantly reduce preparation time. Even though neurotutors open up a promising new possibility for improving the individualized learning process, some challenges still need to be addressed. This will be outlined in the following section.

3.2 Current Challenges in the Development of robust Neurotutors

Most currently available BCIs try to identify specific mental states or processes, as for example processes related to action and perception (e.g., motor imagery; Lotte et al., 2010), cognitive processes (e.g., mental workload; Gerjets, Walter, Rosenstiel, Bogdan, & Zander, 2014) or affective processes (e.g., affective valence; Mühl, Heylen, & Nijholt, 2014). However, systems that only focus on such process-specific user states typically do not account for external influences that might arise during the use of such systems. In line with this reasoning, previous work suggests that spontaneously occurring declines in BCI reliability could be related to modulations in unaccounted mental processes and states (Zander & Jatzev, 2012; Myrden & Chau, 2015). Such performance declines could partly be due to the fact that most currently used classification algorithms assume stationary signal properties, which are derived from a single calibration session. Therefore, these algorithms are incapable to track shifts in the feature space (i.e., changes in statistical signal properties used for BCI classification called non-stationarities). Such classification approaches require the

user's to adapt to unexpected or erroneous behaviors of the system, for instance by trying to produce brain signals similar to those used in the calibration session (Krauledat, Tangermann, Blankertz, & Müller, 2006).

Newly developed, adaptive classification approaches try to reduce this problem using methods from machine learning (Faller, Vidaurre, Solis Escalante, Neuper, & Scherer, 2012; Shenoy, Krauledat, Blankertz, Rao, & Müller, 2006). While such approaches appear to work well in controlled lab environments, we assume that most natural environments for BCI applications (e.g., classroom or workplace settings; cf. Gerjets et al. 2014) contain a multitude of different factors that potentially induce and influence various different mental states and processes. Thus, such application scenarios might easily evoke a variability in different mental states that are not directly related to the core mental state or process that is in the focus of a certain BCI application. Consequently, such factors might induce changes in EEG signals that are picked up by the system, but cannot be interpreted correctly due to the limited framework of the system, which usually only consists of a single mental state or process.

One way to avoid the resulting decline in BCI performance might be the introduction of more realistic frameworks. Such extended frameworks should account for several aspects of the user state, also including potential interactions between different components of the ongoing user state. Up to now, however, only a few studies were even investigating the influence of interactions between different types of mental states and their impact on the classification accuracy of BCI systems. Mühl, Jeunet, and Lotte (2014), for example, tried to estimate mental workload during social stress using features from the frequency domain as well as the time domain. The authors were able to transfer their classification methods across affective contexts, but only with diminished performance. To the best of my knowledge, nobody has yet attempted to formulate a framework that tries to investigate classification problems in BCI systems that involve multiple components of complex user states. While there remain other challenges like accessible hardware, reduction of calibration time and reduction of movement related artifacts, this thesis will focus on the description and measurement of complex user states.

4. Describing Complex User States

User states in the context of BCIs are usually represented as values on a one-dimensional continuum, focusing for example on WML (Gerjets, Walter, Rosenstiel, Bogdan, & Zander, 2014) or affective valence (Mühl, Heylen, & Nijholt, 2014). Such simple representations cannot capture the full complexity of an ongoing user state and therefore might limit the potential for adaption. More advanced frameworks already use two dimensions to reflect cognitive as well as affective aspects of a given user state. For instance, the framework proposed by Fairclough (2009) uses the task engagement dimension to reflect cognitive aspects and the distress dimension to reflect affective aspects (see **Figure 6** for an illustration). These two dimensions form quadrants which can be related to specific user states. Low task engagement and low distress could be related to boredom, which might arise during simple monotonous tasks like vocabulary learning. Low task engagement and high distress could be related to experiences of frustration, which might arise if the task at hand is too difficult. High distress and high task engagement could be related to stress, which might arise during challenging tasks that are straining the limits of an individual capacities. High task engagement and low distress could be related to the experience of flow, which represents a desirable state of optimal performance (Csiksczentmihalyi, Kolo, & Baur, 2004).
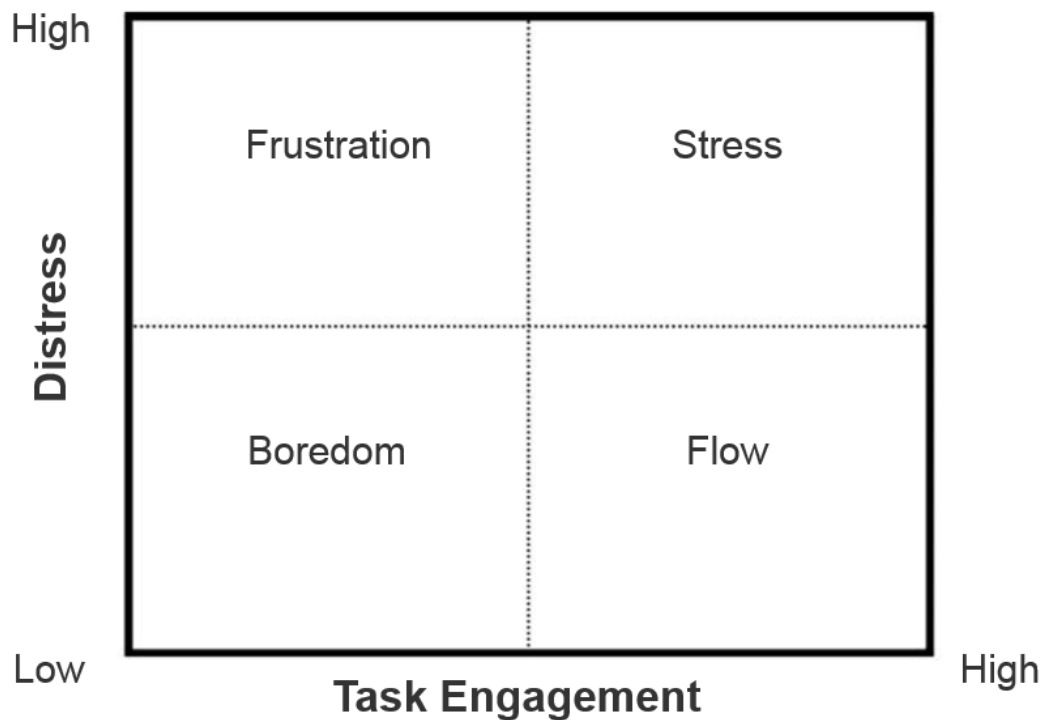
Figure 6. An example for a two-dimensional framework (adapted from Fairclough, 2009).

While this framework seems useful to illustrate the necessity to account for different aspects of a complex user state, it uses higher level concepts which currently have no direct representation in EEG measures. Furthermore, it also does not account for factors related to the type of learning material used as well as factors related to the interaction with the system, like keyboard input.

While recent research (e.g., Brandl, Frølich, Höhne, Müller, & Samek, 2016) has begun to understand the importance of context factors for BCI performance, I do not know of any study that tried to formulate a framework that allows to address this issue in more detail. In this dissertation, I want to propose a framework that accounts for three basic aspects of a given complex user state and try to investigate them as they are currently addressed in BCI research (see **Figure 7** for an illustration of the proposed framework). Furthermore, I collected empirical evidence to test whether the components of the theoretical framework can be separated using the EEG.
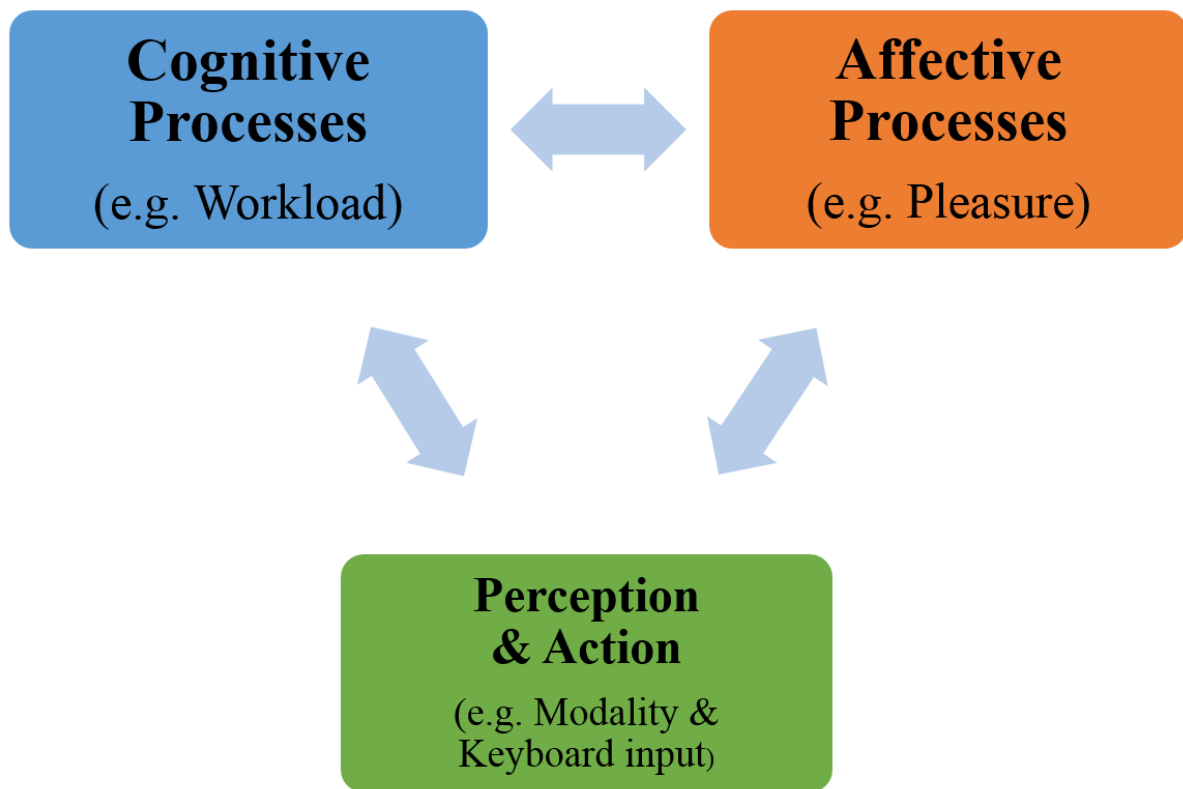
Figure 7. The proposed framework with three components.

The three components of the proposed framework relate to three basic aspects of an ongoing user state, namely cognitive processes, affective processes and perception and action.

The *cognitive processes component* of my framework reflects cognitive aspects of the current user state. Specifically, this component will refer to the detection of WML. WML refers to the cognitive load imposed onto a user due to the complexity of the cognitive demands involved in task processing.

The *affective processes component* relates to aspects of the user state that are associated with affective states, like emotions. Affective states have long been neglected in many research fields (e.g., economics), but recently have received increased attention (Tao & Tan, 2005), especially since approaches focusing on cognitions alone have turned out to be insufficient to explain many aspects of the behavior and experiences of users (Steel & König, 2006). In the field of economics, research has shown that humans generally act less rational than generally assumed (Lopes, 1994). Some researchers even go so far to conclude that the assumption of

rationality is violated by the behavior of every single person (Albanese, 1987). In this thesis this component will focus on the detection of affective valence and dominance.

The *perception and action component* refers to aspects directly related to perceptual interaction with a system like a neurotutor. This component should try to account, for instance for the modality of the presented learning material (e.g., text or videos) as well as the input modality (e.g., responding to questions via keyboard or mouse). Most available active and reactive BCIs are used in a context where direct control of an application via brain activity is required, therefore such BCIs primarily focus on user states that are part of the perception and action component. Due to the fact that only visual material was used in this dissertation this component will mainly account for the motor execution response that occurs due to a keyboard press.

The framework with its emphasis on the interaction of different mental states aims at visualizing complex user states one can assume to arise in real-life environments. Previous research has already shown that affective processes can have an effect on cognitive processes (Fraser et al., 2012). However, before such interactions and possible interferences can be addressed in detail, a framework is needed. To the best of my knowledge nobody has yet attempted to evoke complex user states with the intention to separate different components with neural signatures used in cognitive neuroscience. We assume that cognitive processes and affective processes are most important for educational settings, therefore the main focus of the studies of this dissertation is on these two components as well as their potential interactions.

4.1 Potential Interactions between Working Memory Load and Affective Valence

While different frameworks for the interaction between cognitive and affective processes exist (Power & Dalgleish, 2015), I will shortly mention two that are of specific interest in this context. The theory of hot and cold cognition (Zelazo & Muller, 2002), the capacity model from Ellis and Ashbrook (1989) and the effect of attentional capture (Hodsoll, Viding, & Lavie, 2011).

The theory of hot and cold cognition is based on concepts related to executive function (Zelazo & Muller, 2002). On one hand, cold cognition is based on rational thinking and critical analysis (Roiser & Sahakian, 2013), on the other hand, hot cognition refers to circumstances where a person´s thinking is influenced by their affective state (Brand, 1986). Hot cognition seems able to overpower cold cognition in certain situations and has been shown to impair decision making (Huijbregts, Warren, De Sonneville, & Swaab-Barneveld, 2008).

The capacity model from Ellis and Ashbrook (1989) assumes that there is a limited pool of attentional resources that can be used to complete specific tasks. Affective states can have an influence on the allocation of available attentional resources, drawing from the pool of limited attentional resources and by that potentially impairing task performance.

Attentional capture refers to the automatic diversion of attention toward strongly valenced stimuli (Huang, Baddeley, & Young, 2008). This effect can reduce attentional resources necessary to perform the primary task and should impair task performance especially under difficult conditions.

However, to the best of my knowledge, these theories have not been validated by neuroscience studies that try to infer different components of complex user states for automatic user state detection. Therefore it is unclear in which contexts one can expect interactions between cognitive and affective aspects of a user state in the EEG. Before we look deeper into each of the three components, we have to address the paradigms used to induce complex user states that contain cognitive and affective aspects.

4.2 The Paradigms used to induce User States with cognitive and affective Components

The following sections will shortly describe the two paradigms used to for the studies in this thesis. For a more detailed description see the manuscripts in **section 6**.

4.2.1 The Emoback Paradigm

The first two studies focused on the separation of cognitive processes and affective processes, the two main components of the proposed framework. This was done using a highly controlled task combination of a standard WML induction (N-back task) with a standard affect induction (affective picture stimuli taken from the IAPS database). This task combination is called the emoback paradigm, which has been used to induce different mixtures of WML and affective valence.

The first emoback study used established measures from cognitive neuroscience to separate WML from affective valence. The induction of complex user states was used to assess the context-sensitivity of EEG measures commonly used to infer certain mental states. Furthermore, the induction of differently valenced emotions allowed to check for potential influences of affective processes on cognitive processes.

The second emoback study focused on the automatic detection (classification) of WML as well as affective valence elicited through the emoback paradigm. The emoback task allowed to investigate the potential influence of affective processes on the automatic detection of cognitive processes and vice versa.

4.2.2 The Loss-of-Control Paradigm

The loss-of-control (LOC) paradigm tries to simulate a simplified learning scenario that elicits a complex user state, namely the feeling of perceived LOC. LOC is assumed to contain all three components of the proposed framework during natural human-machine interaction. It seems therefore ideally suited to investigate if the proposed framework can be applied in a less controlled, but more realistic setting. Furthermore, previous work already investigated the influence of LOC on the motor response (Zander & Jatzev, 2012). A related study (briefly summarized in **section 6.3.2**) was focusing on the detection of WML in the LOC data. The LOC study focused on the identification of affective processes and the separation from other components from the complex user state of LOC. Using these related findings all three components of the proposed framework could be addressed.

4.3 Working Memory Load

Our capacity to activate information from long-term memory, process and integrate it using attentional resources is limited. The limits of our capacity to process information is related to the experience of cognitive load (Ellis & Ashbrook, 1989; Plass, Moreno, & Brünken, 2010; Sweller & Sweller, 1994; Wickens, 2008). Different limits of cognitive load or working memory (WM) capacity are related to performance in real-world task settings, which makes it an important concept for educational settings (Cowan, 2014; St. Clair-Thompson & Gathercole, 2006). The next section will shortly describe two relevant WM models.

4.3.1 Theoretical Frameworks for Working Memory Load

Some widely used concepts and theoretical frameworks in the context of WML are the Multi-Component Model (MCM) and Executive Functions (EFs). They will be briefly outlined in the following sections.

4.3.1.1 The Multi-Component Model

The MCM is one widely used theoretical frameworks in the field of WM research and comprises of four components (Baddeley, 2012). The central executive component is related to attention and responsible for the control and regulation of cognitive processes. It is usually seen as the location of EFs that are required to solve different tasks. All other sub-systems are conceptualized as short term memory structures with limited capacity. The visuospatial-sketchpad represents the component where visual information is stored and manipulated. The episodic buffer represents the component used to integrate information from the different sub-systems and link it with chronological information. The phonological loop represents the component that is responsible for the storage and manipulation of phonological information and is therefore mainly engaged in language processing (see **Figure 8** for a visualization).
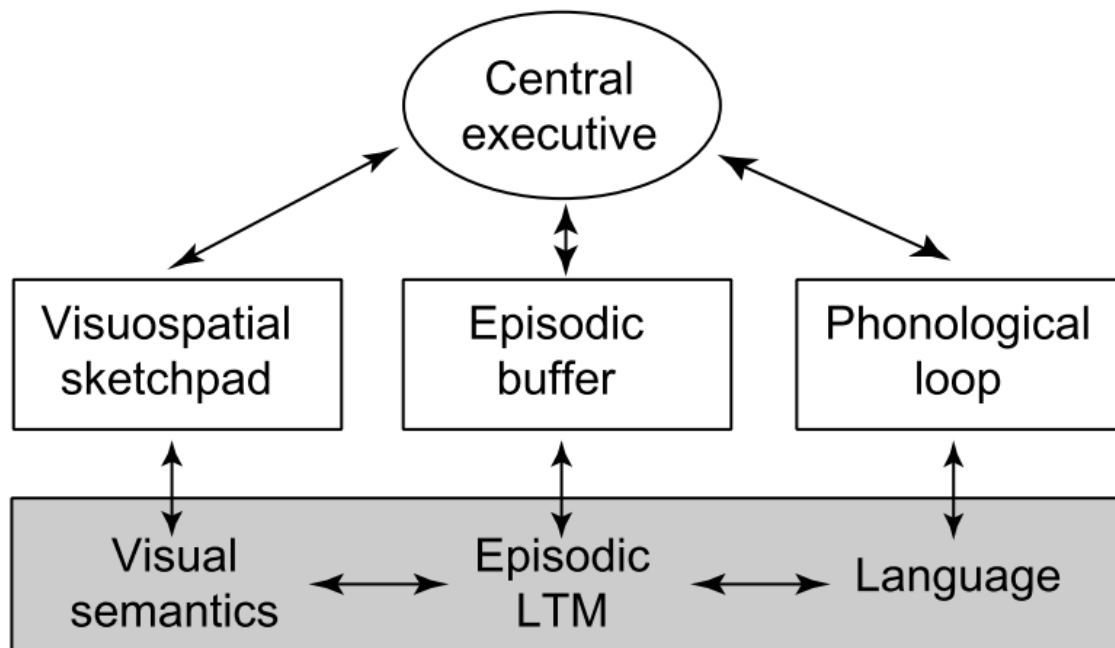
Figure 8. The multi-component model (Baddeley, 2000).

4.3.1.2 Executive Functions

EFs are a central aspect of most theoretical frameworks used in WM research and describe a set of cognitive processes required to solve complex tasks (Diamond, 2013). They range from rather basic cognitive functions, as for example updating, shifting and inhibition (Miyake et al., 2000), called core EFs, to complex, higher order EFs like planning (Jurado & Rosselli, 2007). Most WM processes are closely related to the core EF of updating, which has been associated with the continuous monitoring and manipulation of content in WM. Shifting refers to a change in attentional direction toward relevant information and inhibition refers to the suppression of information that is no longer relevant for the completion of the current task. All these EFs are important for processes related to controlled attention, which is assumed to represent the general factor that includes all core EFs and therefore represents the main limiting factor for WM capacity (Engle, 2002). Interestingly, the relation of WM and EFs seems to depend on the focus of a specific theoretical framework. While the MCM regards EFs as part of WM, attention focused theories consider WM as one EF from many (Diamond, 2013).

4.3.2 Induction of Working Memory Load – The N-back Task

WML is usually modulated via two different approaches. Via changes in task difficulty or via limits of the time made available to complete a certain task (Galy, Cariou, & Mélan, 2012). There are many different tasks used to induce WML (Wilhelm, Hildebrandt, & Oberauer, 2013). One, widely used, way to induce WML in neurophysiological studies is via the N-back task (Jaeggi, Buschkuehl, Perrig, & Meier, 2010; Kirchner, 1958; Scharinger, Soutschek, Schubert, & Gerjets, 2015). The N-back task is a continuous performance task where subjects are presented with a series of stimuli for which they have to indicate via key-press whether the current stimulus is identical to the stimulus presented N-steps before, or not. Hence, the load factor 'N' allows to adjust the difficulty of the task, thereby manipulating WML. The zero-back condition can be seen as a pure matching task, since the target stimuli is already known from the beginning of the task. However, all other load conditions require EFs, mostly updating, since the stimuli list kept in short term memory changes with each newly presented stimulus. Furthermore, the N-back task seems to require inhibition as well, since reactions with regard to the old stimuli have to be suppressed. One can even argue that task switching plays a role in the N-back task. During presentation of the current stimulus, this stimulus has to be matched with stimuli in short term memory, which can be seen as a matching task. After the subjects responded toward the current stimulus, the task switches from matching to updating and inhibition. However, a fine graded discrimination based on this categorization does not seem feasible using established EEG measures at this moment and current approaches only allow for a general estimate of WL load.

4.3.3 EEG Features related to Working Memory Load

The following sections will describe important features for the automatic detection of WML in the EEG.

4.3.3.1 Working Memory Load Features in the Time-Domain – The P300

The P300 represents a positive deflection of the ERP curve which peaks between 250 to 500 ms after stimulus onset and is most prominent over parietal areas (Polich, 2007). The P300 has been linked to different processes relevant for WM tasks like context-updating or attentional factors (e.g., within an oddball paradigm where a new

stimulus creates a novel context and therefore draws more attention; Kappenman & Luck, 2012). Furthermore, a study by Watter, Geffen, and Geffen (2001) found that increased load levels in the N-back task resulted in decreased P300 amplitudes. Still, since approaches based on the time-domain cannot provide continuous measurements of WML (see **section 2.1.1**), this thesis will focus on features from the frequency-domain.

4.3.3.2 Working Memory Load Features in the Frequency-Domain

The automatic detection of WML, necessary for the development of a neurotutor, has already been widely studied in brain-based human-machine interaction research (Berka et al., 2007; Kohlmorgen et al., 2007; Walter, Schmidt, Rosenstiel, Bogdan, & Gerjets, 2013). The most widely used frequency features in this context are modulations in frontal theta activity and parietal alpha activity.

4.3.3.2.1 Frontal Theta Activity

The central attentional control system responsible for EFs is assumed to be mainly located in frontal regions of the brain like the dorsolateral prefrontal cortex. Cognitive processes used to integrate different sources and types of information located in this area are usually reflected in changes of theta band activity (Klimesch, Freunberger, Sauseng, & Gruber, 2008; Mizuhara & Yamaguchi, 2007). Furthermore, theta band activity also seems relevant for memory encoding and may be used to infer individual limits in short-term memory (Jacobs, Hwang, Curran, & Kahana, 2006; Kaminski, Brzezicka, & Wrobel, 2011). Increases in frontal theta band activity are associated with higher WM demands which has been repeatedly found using various different experimental paradigms (e.g., Gruber, Tsivilis, Giabbiconi, & Müller, 2008; Jensen & Tesche, 2002; Osipova et al., 2006).

4.3.3.2.2 Parietal Alpha Activity

The content of different short term memory components is assumed to be maintained in parietal brain areas like the intraparietal sulcus (Klingberg, 2009; Scharinger et al., 2015). Alpha band activity in parietal areas has been found to change in relation to cognitive demands (Klimesch, 1996). More specifically, several studies using a wide range of different stimuli have found that increases in WML are usually reflected in decreased parietal alpha band activity (Gevins & Smith, 2000; Krause et al., 2000; McEvoy, Pellouchoud, Smith, & Gevins, 2001). The function of alpha band activity

could be explained via cortical idling which is seen as the deactivation of a global rest-network due to activation of specific brain areas required to complete a specific task (Pfurtscheller, Stancák, & Neuper, 1996). Hence, alpha band activity is often seen as a counterpart to theta band activity, showing the opposite activation pattern (Klimesch, 1999).

4.3.3.3 Source Space Analysis in the Context of Working Memory Load
Currently, identifying the exact localization of all cortical functions related to WML is not feasible (e.g., the episodic buffer is resulting from the interplay of different components). However, there exists some agreement on the approximate localization of main WM components that are commonly associated with specific sub-components of the MCM (Baddeley, 2003; see **Figure 9**).
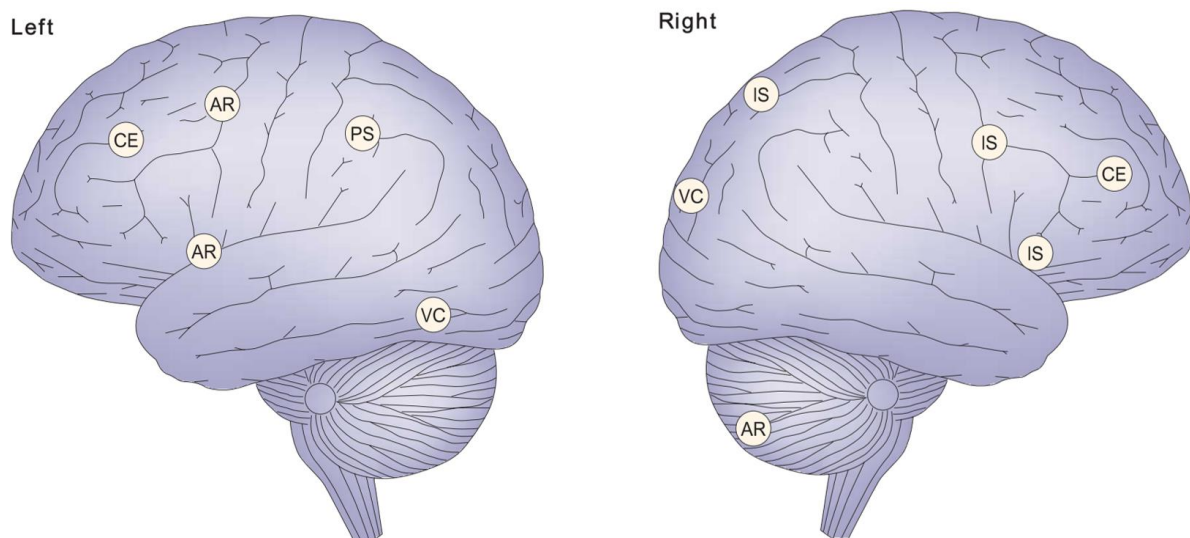
Figure 9. Important components involved in working memory and their approximate locations in the human brain (modified from Baddeley, 2003).

Processes related to the central executive (CE) are usually associated with activity of the dorsal frontal lobe. The left temporal and parietal lobe is mainly related to processes of the phonological loop (articulatory rehearsal [AR] and phonological store [PS]). The right hemisphere is more prevalent for processes related to the visuospatial sketchbook (inner scribe [IS] and visual cache [VC]). Since most WM related studies focus on the presentation of visual material (located in parietal and

occipital areas) and activity in frontal and parietal areas is accessible via the EEG (in contrast to structures located deeper inside the brain), these areas are targeted by most EEG-based WM studies. The rather specific location of functions related to WM makes them ideal candidates for EEG-based source localization methods.

4.4 Affective Valence and Dominance

Affective processes play a major role in human life. Our daily economic decisions are influenced by emotions evoked by marketing campaigns, affective disorders like depression and anxiety influence our health and well-being and, most important for educational settings, achievement emotions like enjoyment and boredom can have a strong impact on learning outcomes (D'Mello et al., 2007; Pekrun, 2006).

Important affective concepts like attitudes, moods and emotions can be distinguished along a temporal axis (Broek et al., 2010; Gross, 2010). Attitudes are beliefs about the goodness or badness of something or someone and therefore relatively stable over time. They can bias how a person will feel, think and behave towards an object or a person (Frijda, 1994). Moods are less stable over time, and in contrast to attitudes, usually do not focus on specific objects. Furthermore, moods seem to mainly influence cognitive processes (Siemer, 2005). Finally, emotions exhibit the fastest response pattern from all affective processes, making it an ideal candidate for the investigation with the EEG. Giving exact definitions on the term 'emotion' can be difficult (Hendelman, 2005). A review by Kleinginna and Kleinginna (1981) represents a good exemplification of various concepts and definitions related to emotions. The basic elements commonly agreed upon are that emotions include physiological change (e.g., induced by basic drives like thirst or sexual behavior), expressive behavior (e.g., facial expressions) and corresponding changes in mental states (e.g., psychological reaction toward a stimulus, called subjective experience; Luay & Revett, 2011). Consequently, emotions can be seen as fast reactions in experiential, behavioral as well as physiological response channels to the perception of a specific, relevant (internal or external) stimulus (Levenson, 1999). **Figure 10** illustrates the key terms used in affective science.
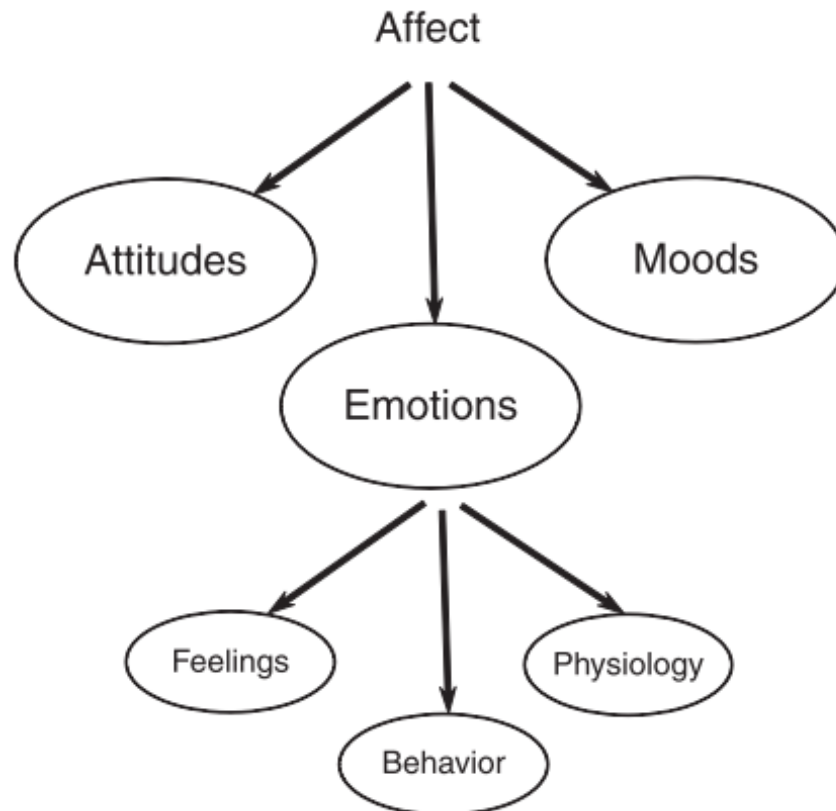
Figure 10. Key terms in affective science (Gross, 2010).

4.4.1 Theoretical Frameworks for Emotions

There are various theoretical frameworks that describe emotions. The two most widely used in affective neuroscience are discrete or basic emotion models and dimensional emotion models.

4.4.1.1 Discrete Emotion Models

The oldest theoretical frameworks used to describe emotions in the context of affective neuroscience are the basic or discrete emotion models (see Dalgleish, 2004 for a short review). These models are rooted in the language of daily life and use discrete categories to describe different emotions (Tomkins, 2008). The most popular emotion categories are happiness, sadness, fear, anger, disgust, and surprise for which cross-cultural research has provided evidence to their universal validity (Ekman et al., 1987). The main advantage of this theoretical framework is the labeling scheme. It is based on daily experiences that can be observed (e.g., happiness is usually accompanied by a smile) and this labeling scheme is therefore very intuitive. Though, due to the focus on strong, prototypical emotion categories, discrete models are often

unable to convey all the different nuances of emotional experiences occurring in our daily lives. This led to the development of dimensional emotion models.

4.4.1.2 Dimensional Emotion Models

Dimensional models try to explain emotional experiences along a small number of latent dimensions (Zeng, Pantic, Roisman, & Huang, 2009). Accordingly, such frameworks allow to represent each emotion as point in a coordinate system, even if there is no label or expression available to define it. Additionally, such dimensional models allow to represent the intensity of an emotion. A point in the middle of such a coordinate system would represent a low intensity or neutral emotion, while points at the outer border would represent emotions with strong intensity.

4.4.1.2.1 The Circumplex Model

The most widely used dimensional framework for the description of emotions is the circumplex model, which uses two dimensions to describe the affective space (Russell, 1980). The first dimension is the valence dimension which relates to the pleasantness or unpleasantness of an emotional experience. This valence or evaluation dimension measures how an individual feels and ranges from extreme happiness (ecstasy) to complete sadness. The second dimension is the arousal or activation dimension which is associated with the responsiveness of an individual. It indicates if a person is likely to take action (Zeng et al., 2009). The arousal dimension can range from total relaxation (or sleep) to frenzied excitation. A specific emotion can be seen as a mixture of these two dimensions (e.g. excitation has high arousal and positive valence). The angle from the middle of the space (= the intersection of both dimensional axes) defines the quality of the experienced emotion and the distance from the middle can be used to measure the intensity of the experienced emotion (see **Figure 11** for an illustration). Different emotions are arranged along the two-dimensions forming a circle, hence the name of the model - circumplex.
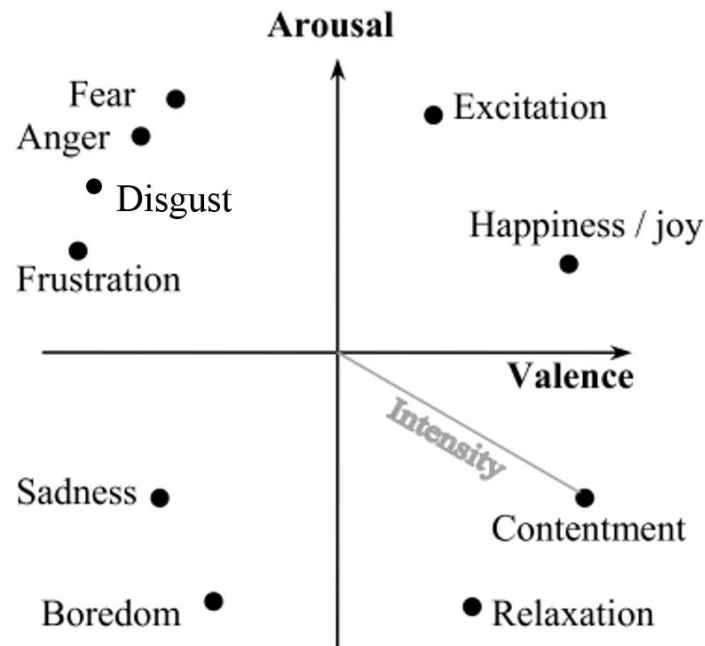
Figure 11. The two-dimensional affective space of the circumplex model (based on Russell (1980) and modified from Chanel (2009).

While Russell, Weiss, and Mendelsohn (1989) argued that the emotional dimensions of valence and arousal are independent, Lang, Bradley, and Cuthbert (2005) observed that when people judge their own emotion while watching affective pictures, their judgements tend to follow a distorted 'U' or banana shape (see **Figure 12**). This distribution of self-judgements is due to the difficulty of eliciting emotions with high valence and low arousal as well as emotions with high arousal and low valence.
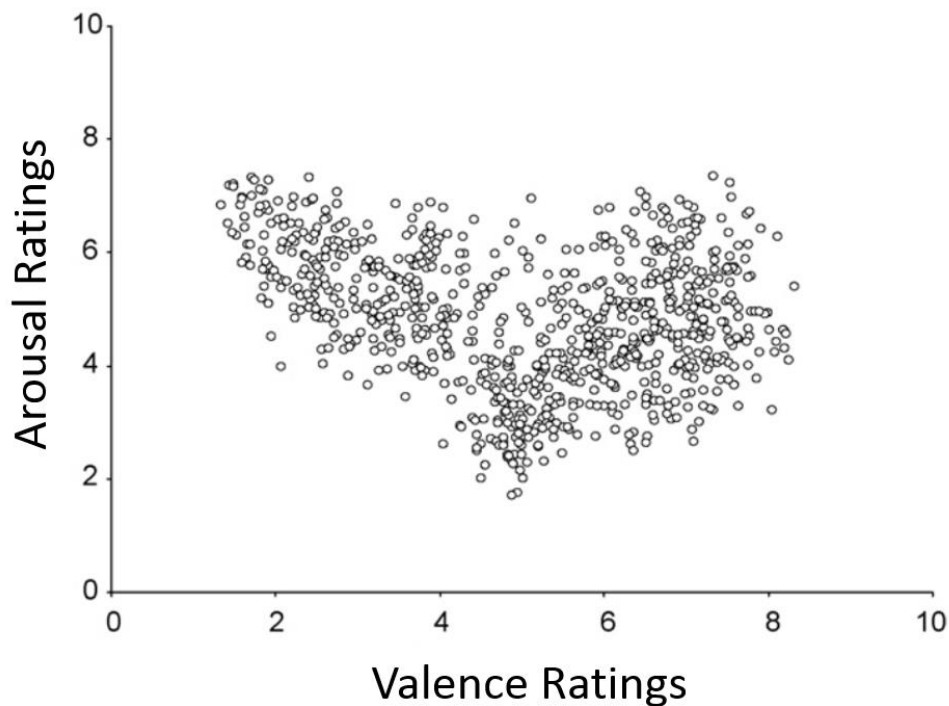
Figure 12. Self-assessments using the valence-arousal framework. (modified from Olofsson, Nordin, Sequeira, & Polich, 2008)

Since the circumplex model originated from the analysis of emotional expressions it is possible to project emotional labels from discrete emotional models onto the continuous valence-arousal space (Russell, 1980). Unfortunately, this can result in similar points for quite distinct emotions. The most important case is the distinction between anger and fear, since both have negative valence and high arousal. This is an important distinction for neurotutors, because test anxiety plays a major role in the educational context.

4.4.1.2.2 Test Anxiety as an Example for the Importance of the Distinction of Anger and Fear

Test or achievement anxiety is a prevalent issue in modern societies. Our occupational careers strongly depend on our academic performance which is the principle of the meritocratic societies we currently live in. Educational careers depend on individual achievements, which leaves only small room for error if one wants to avoid failure. Hence, the anxiety of failure is one of the most frequently experienced emotions in achievement settings like the classroom (Pekrun, Goetz, Titz, & Perry, 2002). Test anxiety can have detrimental effects on performance during difficult tasks, motivation

and personal health and well-being (Hembree, 1988). Its relevance for educational outcomes has been recognized early, with scientific investigations dating back to 1936 (Stengel, 1936). Over the years a multitude of different theories have been developed and more than thousand empirical studies have been conducted on the topic of test anxiety (Zeidner, 1998). It is thus, the most intensively studied emotion in educational and achievement settings. This highlights the relevance of test anxiety for educational settings. To successfully detect test anxiety with a neurotutor, one needs to distinguish it from other affective concepts like anger. This is something which might be achieved using the pleasure-arousal-dominance model proposed by Mehrabian (1996).

4.4.1.2.3 The Pleasure-Arousal-Dominance Model

The distinction between fear or anxiety on one hand and anger on the other hand can be established with the introduction of a third factor or dimension, the so-called dominance or control dimension. First evidence for a three factor solution was already discovered in 1977 by Russell and Mehrabian. The dominance dimension can be defined by 'feelings of control and influence over everyday situations, events, and relationships versus feelings of being controlled and influenced by circumstances and others (Mehrabian, 1994). When an individual is anxious or afraid, he (or she) does not have control over the situation, while during the experience of anger there usually is the experience of a certain amount of control. The dominance can also be used to distinguish self-confidence (moderate dominance) from arrogance (high dominance) and awe (low dominance). While self-confidence and awe are desired states in an educational setting, arrogance could lead to misjudgments concerning ones abilities.

This extension of the two-dimensional circumplex model led to the construction of the pleasure-arousal-dominance model (PAD), where pleasure is seen synonymous to the valence dimension of the circumplex model (Mehrabian, 1996). To enable easier reading of the text, all future instances related to the pleasure dimension will be termed valence dimension. An illustration of the PAD can be seen in **Figure 13**.
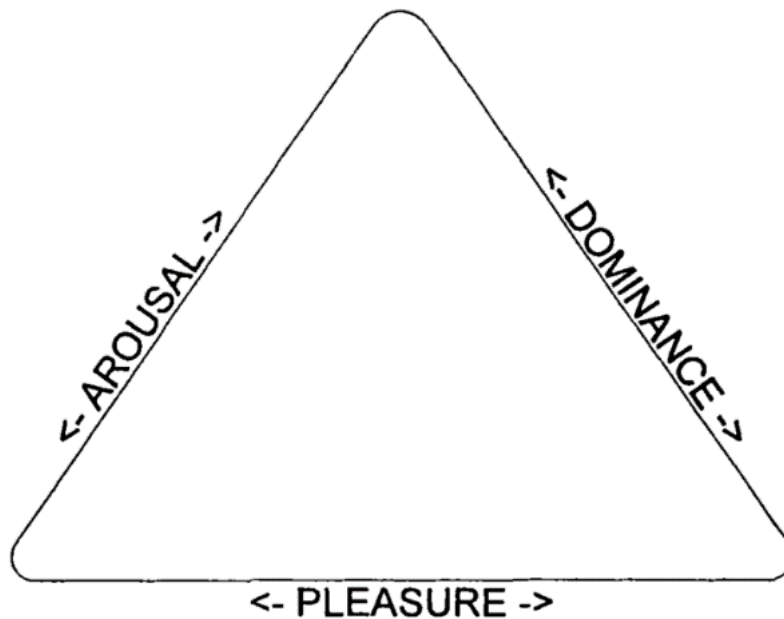
Figure 13. The pleasure-arousal-dominance model (Mehrabian, 1996).

4.4.2 Induction Methods for Emotions

A variety of different emotion induction methods have been used in the past in the field of emotion research. Gerrards-Hesse, Spies, and Hesse, (1994) provide an overview, using five different categories to describe these different approaches. One way to induce emotions is the free mental generation of emotional states via hypnosis or imagination (e.g., Makeig et al., 2011). Another way to induce emotions is the guided mental generation of emotional states via stories, music or movies accompanied by additional instructions on how to enter the target state (e.g., Onton & Makeig, 2009). Furthermore, emotions can be induced via the presentation of affective situations that activate certain needs (e.g., the need for achievement), which is usually induced by manipulation of success or failure in experimental settings (e.g., Jatzev, Zander, Filippis, & Kothe, 2008; see also study in **section 6.3**). Additionally, emotions can be induced via the generation of emotionally relevant physiological states (e.g. via the facial feedback paradigm; Mikhail, El–Ayat, Coan, & Allen, 2013). Finally, one widely used way to induce emotions is via the presentation of emotion inducing material like affective pictures (Lang, Bradley, & Cuthbert, 1997). Affective pictures allow for a fast and efficient induction of affective responses, which is necessary to investigate

emotions with the EEG and were therefore chosen for the first paradigm used in this thesis (see studies in **section 6.1** and **section 0**).

4.4.2.1 The International Affective Picture System

The International Affective Picture System or IAPS is a database which is widely used to induce emotional experiences. It consists of more than eight hundred emotionally laden color pictures which are rated along the three important emotional dimensions of valence, arousal and dominance by a large number of participants using a visual analog rating scale (Lang et al., 2005). This emotion induction is easy to implement and very effective for inducing negative emotions.

4.4.3 EEG Features related to Emotions

The next section will describe EEG features from the time-domain as well as the frequency-domain, used to infer affective user states.

4.4.3.1 EEG Features to infer Affect in the Time-Domain

A significant amount of affective EEG research investigated the impact of emotional processes on ERPs (see Olofsson et al., 2008, for a review). The two most widely used potentials in this context are the P300 and the late positive potential (LPP). The P300 is associated with the orientation toward very salient (or rare) stimuli (Polich, 2007). It has been shown to be influenced by motivation (Kleih, Nijboer, Halder, & Kübler, 2010) and in response to strongly negative stimuli (Briggs & Martin, 2009). The LPP is a positive deflection in response to an emotionally arousing visual stimuli (Schupp et al., 2000). It has been associated with increased perceptive evaluation of emotionally salient stimuli which is reflected in higher activity in posterior visual brain areas (Sabatinelli, Lang, Keil, & Bradley, 2007). While this thesis will focus on the valence and dominance dimension, the LPP is mostly related to the arousal dimension. Additionally, as mentioned in **section 2.1.1**, ERPs might not be well-suited for the use in neurotutors, due to their missing capability to provide continuous measures. I therefore focus on signals from the frequency-domain to infer affective valence in my dissertation.

4.4.3.2 EEG Features to infer Affect in the Frequency-Domain

The most prominent EEG feature in the frequency domain with regard to affective processes is a frontal asymmetry in the alpha range, called Frontal Alpha Asymmetry (FAA; Demaree, Everhart, Youngstrom, & Harrison, 2005). It is assumed that the different lateralization of frontal alpha power during affective processing is due to the different roles the left and the right prefrontal structures play in the processing of affective material. This assumption is reflected in the valence model which proposes that the left hemisphere is specialized for positive emotions and the right hemisphere is active when negative emotions are processed (Mills, 1912, as cited in Dalgleish, 2004). Most studies that used brain lateralization as a measure of affective states or processes focused on the valence dimension (Davidson, 1992; Davidson & Tomarken, 1989; Huang et al., 2012; Lin, Wang, Wu, Jeng, & Chen, 2008; Ramirez & Vamvakousis, 2012; Silberman & Weingartner, 1986). Left frontal activity is seen in this context as a tendency toward positive emotions like happiness (e.g., Ahern & Schwartz, 1985). Yet, several studies showed that emotions with negative valence can also activate the left hemisphere which is reflected in the approach-withdrawal model developed by Schneirla (Schneirla, 1959, as cited in Dalgleish, 2004). This model states that the left hemisphere is active in emotions associated with approach behaviors like engagement or anger and the right hemisphere is related to avoidance behaviors like anxiety. Furthermore, recent evidence indicates that the FAA can be used to measure activity related to emotional responses along the dominance dimension (Demaree et al. 2005). In this particular context left-frontal activation is associated with increased dominance. Interestingly, a study by Reuderink, Mühl, & Poel (2013) found that dependent on the type of alpha band, the FAA responded differently to affective valence and dominance. This reactivity towards affective responses with regard to the valence as well as dominance dimension underscores the potential of the FAA for the EEG-based detection of emotions within the PAD. This thesis will focus on the FAA with the intention to infer affective responses relating to the affective dimensions of valence and dominance.

4.4.3.3 Analysis of the Source Space in the Context of Emotions

While it is hard to pinpoint specific brain regions underlying functional asymmetries in the human brain (Wager, Phan, Liberzon, & Taylor, 2003), first successful attempts using ICA-based approaches for affect detection have already been made. A study by

Onton and Makeig (2009) used an eyes-closed emotion imagination task and was able to reconstruct the valence-arousal space using ICA. Another example is the study by Makeig, Leslie, and Mullen (2011) which associated feelings with musical sounds for a live demonstration of an emotion BCI. Analysis of source signals using an ICA approach showed that different emotion imaginations were related to specific brain dynamics. This indicates that EEG-based source localization can help to further identify the underlying brain dynamics behind established EEG features.

4.5 Perceptual Interaction

The main focus of this thesis is on the separation of cognitive (i.e. WML) and affective processes (affective valence and dominance). Nevertheless, use of a neurotutor also requires active interaction between the user and the system which is usually achieved via mouse or keyboard input. Additionally, responses in the EEG are depending on the stimulus modality used in the system. A study by Mühl et al. (2011), for example, used visual, auditory and audio-visual stimuli for affect induction and found opposing patterns for visual and auditory affective stimulation. The authors concluded that a classifier trained on stimuli from a specific modality will be limited in its capacity for generalization toward other stimulus types. Additionally, one can take the case of the induction of WML using visual stimuli, as is the case in this thesis. Increases in WML are accompanied by decreased parietal alpha activity. However, visual stimulation due to the presentation of complex stimuli also usually results in decreased occipital activity (i.e., alpha-blocking), which could spread to parietal areas via volume conduction. Accordingly, to be able to develop robust systems two additional aspects need to be accounted for. Firstly, different stimulus modalities have to be taken into account, which is reflected in the perception component of the proposed framework. Secondly, motor responses related to interaction with the system need to be accounted for. This is reflected in the action component of the proposed framework. EEG responses related to activity of the sensorimotor cortex are reflected in alpha band changes (mu-rhythm, see **section 2.1.2**; Pfurtscheller, Neuper, Andrew, & Edlinger, 1997) in central areas. This could result in overlaps with signals used for the detection of cognitive processes reflected in parietal alpha responses due to changes in WML (see **section 4.3.3.2.2**) and affective processes like FAA responses due to changes in emotional

experience (see **section 4.4.3.2**). Such potential overlaps at the scalp level might be addressed with source localization approaches, but before one can attempt this one needs to account for such aspects of the ongoing user state. I therefore included the perception and action component in the proposed framework.

5 Aim of the Dissertation Project

In this dissertation I want to investigate the potential of the EEG for the detection of complex user states that consist of mixtures of different cognitive as well as affective processes. EEG measures commonly used to infer these aspects will be tested for robustness in paradigms that evoke complex user states. To be more specific, I want to investigate if WML and affective valence or dominance can be identified using EEG measures usually associated with them under conditions that more closely resemble a realistic situation. Futhermore, said states should be automatically detectable using methods from a BCI framework.

5.1 Research Questions

Based on the assumption that natural environments elicit complex user states that consist of a mixture of different basic components I proposed a framework which should help to investigate such complex user states with the intent to automatically detect them via the EEG. Accordingly, I formulated the following research questions (RQs).

RQ1: *Can established EEG measures from the frequency-domain be used to separate WML from affective valence or dominance using paradigms that evoke complex user states?*

I will try to answer this RQ in two steps. Firstly, I used a controlled study (emoback) that allowed to independently modulate WML and affective valence. Secondly, I used a paradigm that simulates a simple learning scenario which evoked the complex user state of perceived LOC. This second paradigm is less controlled, but much closer to the target scenario.

RQ1a: *Can EEG-based source localization methods be used to identify cortical sources of the signals recorded at the electrode level?*

I applied ICA-based source localization methods to identify cortical sources of EEG responses that are recorded at the electrode level in the LOC study.

RQ1b: *Are affective processes related to the valence dimension separable from affective processes related to the dominance dimension?*

I tried to separate affective responses related to the valence dimension from affective responses related to the dominance dimension in the LOC study. This was done via a separate analysis of the lower and the upper alpha band.

RQ2: *Are there interactions between the EEG measures used to infer WML and affective processes related to the valence or the dominance dimension?*

The second research question focuses on potential interactions between WML and affective valence, which will be addressed with the first paradigm (emoback). Additionally, I want to investigate if the proposed framework can help to integrate the findings from multiple studies that relate to different aspects of a complex user state. This will be done using the results from three different studies that all investigated the same LOC dataset.

RQ3: *Are WML and affective valence automatically detectable (classifiable) with accuracies significantly higher as chance in paradigms that evoke complex user states?*

Since measuring different aspects of a complex user state does not necessarily mean that those aspects can be used during human-machine interaction, the third RQ addresses the automatic detection of WML and affective valence as components of a complex user state.

6   The three Manuscripts describing the Studies of this Thesis

The next sections will contain all three manuscripts used as basis for this thesis and additional results which were not part of the submitted manuscripts.

6.1   Electroencephalography based identification of working memory load and affective valence in an N-back task with emotional stimuli

(Emoback Study 1)

Sebastian Grissmann[1]*, Josef Faller[2], Martin Spüler[3], Thorsten O. Zander[1,4], Augustin Kelava[1,5], Christian Scharinger[6] & Peter Gerjets[1,6]

[1]LEAD Graduate School, University of Tübingen, Tübingen, Germany

[2]Laboratory for Intelligent Imaging and Neural Computing, Columbia University, New York, USA

[3]Wilhelm-Schickard-Institute for Computer Science, University of Tübingen, Tübingen, Germany

[4]Team PhyPA, Biological Psychology and Neuroergonomics, Berlin Institute of Technology, Berlin, Germany

[5]Hector Research Institute of Education Science and Psychology, University of Tübingen, Tübingen, Germany

[6]Leibniz-Institut für Wissensmedien, Tübingen, Germany

Abstract

Up to now, most brain-based measures of the electroencephalogram (EEG) are used in lab environments and only focus on narrow mental states (e.g. WML). However, we assume that outside the lab complex multidimensional mental states are evoked. This could potentially create interference between EEG signatures used for identification of specific mental states. In this study, we investigated interactions between WML and affective valence to reveal potential interferences in EEG measures. To induce changes in WML and affective valence, we used a paradigm which combines an N-back task (for WML manipulation) with a standard method to induce affect (affective pictures taken from the International Affective Picture System or IAPS database). As expected, we found that behavioral performance decreased with increasing workload as well as negative valence, showing that affective valence can have an effect on

cognitive processing. These findings are supported by changes in frontal theta and parietal alpha power, parameters used for measuring of WML in the EEG. However, these EEG measures are influenced by the negative valence condition as well and thereby show interactions between WML and affective valence. Unexpectedly, we did not find any effects for EEG measures typically used for affective valence detection (frontal alpha asymmetry or FAA). Therefore we assume that the FAA measure might not be usable if cognitive workload is induced simultaneously. We conclude that future studies should account for potential context-specifity of EEG measures.

Introduction

Investigating Complex User States with the Electroencephalogram

Most studies using the electroencephalogram to measure mental states focus on very specific states like working memory (Klimesch, 1999) or affective valence (Ahern & Schwartz, 1985), which are investigated in well controlled lab environments. Therefore, the indicators used in such studies might not provide robust measurements outside the lab, since real world environments tend to evoke much more complex and multidimensional mental states that involve different cognitive, emotional and motivational components (Gerjets et al., 2014). Furthermore, many measures used to infer mental states from the ongoing EEG are known to have many-to-many relations, meaning that several physiological variables are associated with multiple psychological elements (Fairclough, 2009). Hence, it is necessary to systematically investigate the relationship between different mental states and the different EEG measures that are widely used in neuroscientific studies, to investigate if such measures can be used outside the lab. In this paper we systematically investigate the relation between two types of mental states, namely WML and affective valence, to study the interaction of the brain responses typically associated to these mental states

Working Memory Load

There are many different ways to induce mental states characterized by high levels of WML (Wilhelm et al., 2013). One way to induce WML that has been widely used in the context of cognitive neuroscience is the N-back task (Kirchner, 1958; Scharinger et al., 2015). The N-back task is a continuous performance task where subjects are presented with a series of stimuli and have to indicate whether the current stimulus is

identical to the stimulus presented N-steps before, or nor. Hence, the load factor 'N' allows to adjust the difficulty of the task, thereby manipulating WML.

WML is commonly defined as the interplay of controlled attentional processes and short term memory structures that handle different representational codes via various temporal storage components (Baddeley, 2003, 2012). The central attentional control system is assumed to be mainly located in frontal regions of the brain like the dorsolateral prefrontal cortex, while content in short term memory is thought to be maintained via parietal brain areas like the intraparietal sulcus (Klingberg, 2009; Scharinger et al., 2015). Accordingly, increases in WML usually result in increased frontal theta activity as well as decreased parietal alpha activity (Gevins et al., 1995; Klimesch, 1999; Smith & Gevins, 2005). However, previous research has shown that other mental states can also have an effect on measures used for workload detection. Roy et al. (Roy, Bonnet, Charbonnier, & Campagne, 2013), for example, induced fatigue and found that with increasing time on task the discriminability of WML was decreased.

Affective Valence

There are multiple ways to induce affective mental states in lab settings (Gerrards-Hesse et al., 1994). One effective way is the use of affective picture stimuli. The most prominent picture database is the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1997) which comprises a large set of standardized and emotionally evocative color photographs. All stimuli of the database have been rated along the dimensions of valence and arousal as described in the two-dimensional circumplex model of emotion (Russell, 1980).

The valence dimension reflects the pleasantness of a situation and ranges from sadness to happiness. The arousal dimension reflects the responsiveness of the organism and ranges from sleep to frenzied excitement. Several studies have used the EEG to study affective states in the past (Kim, Kim, Oh, & Kim, 2013; Olofsson et al., 2008). The two most widely used EEG measures to infer affective states are the late positive potential (LPP) and the so-called frontal alpha asymmetry or FAA. The LPP is an EEG feature in the time domain and represents a positive deflection in the ERP-curve, reflecting the activity related to the arousal dimension (Schupp et al., 2000). The FAA represents the individual hemispheric contributions which is related to the

affective valence dimension (Ahern & Schwartz, 1985; Huang et al., 2012; Tomarken, Davidson, & Henriques, 1990). Increased right frontal activity is an indicator of a mental state characterized by negative affective valence. Usually, this measure is used during resting (with closed eyes) or passive viewing conditions (Davidson, 1992). Interestingly, previous work has shown that WML can result in lateralized activity as well. For instance, a study by Baldwin and Penaranda (Baldwin & Penaranda, 2012) used several tasks to induce mental workload and found more left hemispheric activity during increased cognitive load. This might result in potential interferences between affective valence and WML in the FAA measure. Analyzing this type of interaction between WML and affective valence is the main goal of this paper.

Investigating Interactions between Working Memory Load and Affective Valence: The Emoback Task

Previous research on the use of neural signatures of mental states has largely ignored the problem of potential interference between WML and affective states. Only one recent study addressed the effect of an affective state on the identification of WML. In this study, Mühl, Jeunet, & Lotte (Mühl, Jeunet, et al., 2014) used an N-back task to manipulate WML while social stress was induced with a stress-induction protocol based on the Trier Social Stress Task (Kirschbaum, Pirke, & Hellhammer, 1993). The authors attempted to detect mental workload during stress using features from the frequency domain as well as the time domain. They concluded that it is possible to transfer classification methods across affective contexts, but only with diminished performance. However, these results are limited to one specific affective context (social stress). Furthermore, the authors were focusing on the identification of WML alone. The current paper wants to extend these findings using a more general affective response and also account for the interaction of cognitive and affective components.

In order to collect suitable data for this objective we used a combination of an N-back task with a standard affect induction, called the emoback task. Interactions between cognitive and affective processes in the EEG have been previously investigated with the affective flanker task (see Alguacil, Tudela, & Ruz, 2013). However, the affective flanker task can be seen as a simple stimulus-response task that only requires perceptual inhibition and therefore does not represent a working memory task. The emoback task does involve memory components and, to our

knowledge, has only been used twice in combination with the EEG. First, a study by MacNamara, Ferri, & Hajcak (2011) used the emoback task with affective pictures as distractors and found that the LPP was modulated not only by affective responses towards emotional pictures, but also by WML. Second, a study by Kopf et al. (2013) used a N-back task with emotional words and recorded data from EEG and fNIRS. They found more errors during the negative condition, especially for high task difficulty. An ERP-analysis also revealed that the LPP is influenced by the difficulty in the working memory task and that this influence is further modulated by affective valence. However, both these studies used the LPP, as feature commonly associated with the arousal dimension. We want to investigate affective responses related to the valence dimension, which allows to make more general discriminations of affective states into positive and negative states. Moreover, the study by MacNamara, Ferri, & Hajcak (2011) only used affective stimuli as distractors, while we want to use a paradigm that inherently activates cognitive as well as affective processes. Finally, while the study by Kopf et al. (2013) used emotional words to induce affective reactions, we assume that affective pictures can elicit stronger affective reactions. We therefore decided to use an emoback task with affective pictures from the IAPS database to investigate interactions between WML and affective valence using frontal theta activity, parietal alpha activity as well as the FAA.

Research Questions and Hypotheses

We investigated the influence of WML and affective valence on behavioral measures as well as the corresponding EEG measures. Furthermore, we investigated potential interactions between EEG measures used to infer WML and EEG measures used to infer affective valence.

More specifically, we investigated potential effects of load levels in the emoback task on accuracies and reaction times, which might also be reflected in the corresponding EEG measures. Additionally, we investigated potential effects of the affective valence inductions in the emoback task on behavioral measures as well as EEG measures.

Beyond the main effects of WML and affective valence on behavioral measures and EEG measures, we also analyzed whether EEG measures used for mental

workload detection are sensitive to different affective contexts and whether EEG measures used to infer affective valence are sensitive to WML.

Methods

Sample

We collected data from 27 female subjects who were recruited from an online database. Female subjects were used in this study because they tend to show stronger reactions toward affective stimuli (Lang, Greenwald, Bradley, & Hamm, 1993) and also exhibit more stable responses (Ahern & Schwartz, 1985). Three subjects had to be removed due to bad data quality. All of the participants were university students, aged above 18 years (mean: 23.0 years; range: 19 to 32 years), right handed and had no blood phobia to avoid extreme responses toward the experimental stimuli. All subjects provided written informed consent and were paid 20 € for participation in the experiment. The study was approved by the ethic committee of the Knowledge Media Research Center Tuebingen.

Recording of Physiological Data

Sixty channels of EEG were recorded using an ActiCHamp amplifier and active Ag/Cl-electrodes (Brainproducts GmbH, Gilching, Germany). Electrodes were placed according to the extended 10-20 system. The electrooculogram (EOG) was recorded with four EEG electrodes located at the left and right canthi as well as above (channel Fp1) and below the left eye. All channels were referenced to channel FCz. Impedances were kept below 10 kΩ. The data were sampled at 1 kHz. For later processing the data was downsampled to 250 Hz. During the recordings subjects were instructed to sit in a relaxed posture to avoid artifact contamination of the data.

Preprocessing

To automatically reject time windows that are contaminated by artifacts, a time window was removed if channel power at more than six channels exceeded five times the standard deviation. After computing independent components using the CUDAICA implementation (Raimondo, Kamienkowski, Sigman, & Fernandez Slezak, 2012) of the Infomax independent component analysis (Bell & Sejnowski, 1995), eye movement artifacts were removed (reduced) using the ADJUST approach (Mognon, Jovicich,

Bruzzone, & Buiatti, 2011) which is implemented as plug-in in the EEGLAB toolbox (Delorme et al., 2011). Finally the EEG signal was bandpass filtered between 1 and 45 Hz.

Stimuli

For the positive, neutral and negative valence conditions we selected 96 stimuli each. The stimuli were taken from the IAPS database (Lang et al., 1997) and selected based on the valence and arousal ratings provided with the database. Valence and arousal ratings are usually confounded, meaning that stimuli with strong (positive or negative) valence ratings usually also have high arousal ratings. To improve discriminability between affective conditions we made sure that stimuli for the positive condition had the highest valence ratings while stimuli for the negative conditions had the lowest valence ratings. Furthermore, stimuli for the neutral condition were selected based on the lowest arousal ratings. All selected stimuli had a quadratic shape and were centrally presented on a standard 23 inch display. To avoid unnecessary eye-movements the size of the stimuli was scaled to fill 60 % of the height of the display.

Study Design and Block Structure

The affective valence conditions were presented in groups of four trial blocks with identical affective valence. The affective valence conditions being either positive, neutral or negative. Load levels of the emoback task were alternated block wise. The load factor of the emoback task had two levels, one-back and two-back. We avoided the use of a zero-back condition, because this would require to repeatedly use the same affective target stimuli. This repetition of the same stimuli might have resulted in affective habituation with regard to the target stimuli, thereby diminishing the affect induction (Leventhal, Martin, Seals, Tapia, & Rehm, 2007). We also avoided a 3-back condition because some studies showed that this load level can lead to task disengagement (Ayaz, Izzetoglu, Bunce, Heiman-Patterson, & Onaral, 2007). Target response hand as well as starting load level of the emoback task were balanced across the subject sample (see **Figure 14**).

Positive Condition				Neutral Condition				Negative Condition			
1-back	2-back	1-back	2-back	1-back	2-back	1-back	2-back	1-back	2-back	1-back	2-back
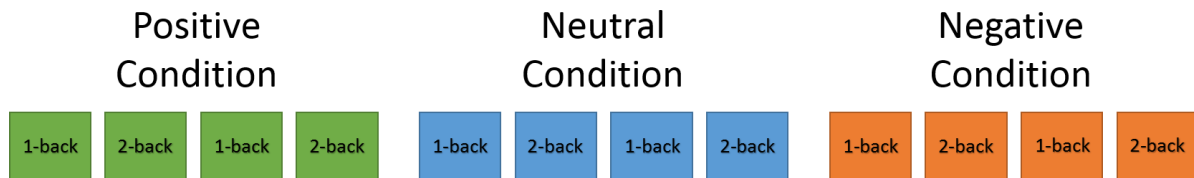
Figure 14. Illustration of emoback study design: The three affective valence conditions were presented grouped and permutated across the whole sample. Each affective condition consisted of four workload blocks, resulting in 12 blocks for each subject. Workload blocks were presented in an alternating fashion and balanced across the whole sample.

All participants performed 12 blocks in total. Each block started with a 10 second baseline where subjects were instructed to relax and visually fixate a centrally presented light grey fixation cross (see **Figure 15**). After the baseline the first trial started with a stimulus presentation. There were 72 trials in one block, consisting of 24 target stimuli and 48 distractors. The 24 target stimuli were randomly selected for each subject from 96 unique stimuli (four blocks with 24 target stimuli per affective condition). Targets and distractors were sometimes interleaved in the 2-back condition, meaning that two target stimuli could appear right after each other. Here is an example: Distractor (e.g. table), distractor (e.g. chair), target (table) and target (chair). After the last trial, another baseline phase, also with a duration of 10 seconds, was recorded. Between blocks there were short brakes between one and three minutes.
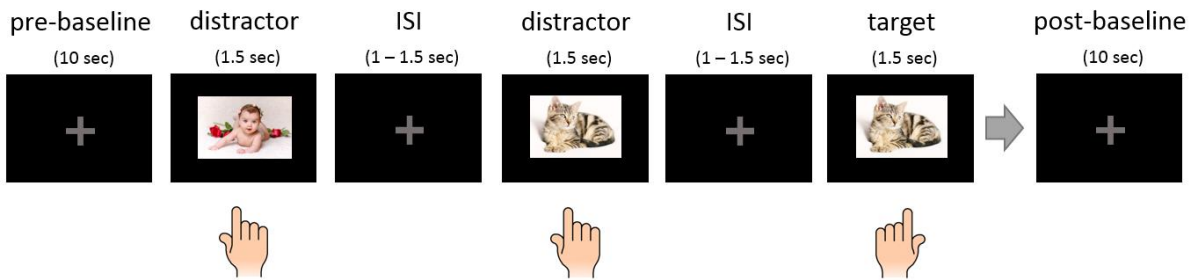
1-back condition
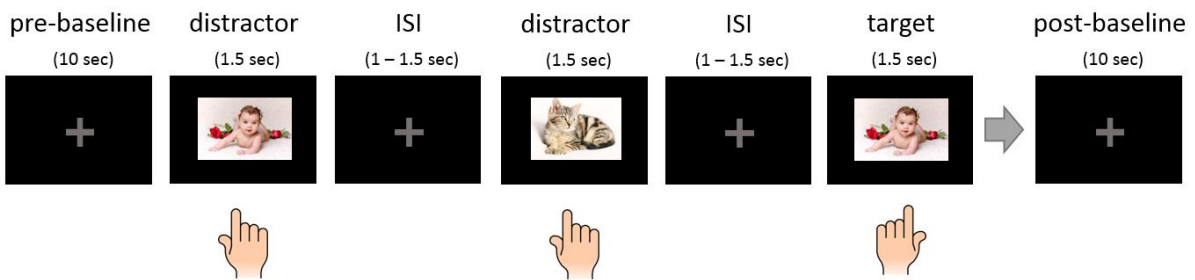
2-back condition

Figure 15. Schematic representation of emoback block structure. Top: 1-back condition. Bottom: 2-back condition. Hand symbols indicate key press. Left hand press was used for target stimuli and right hand press for distractors in this example Pictures in this illustration are taken from https://pixabay.com/ and only serve as an example.

Trial Structure

Each trial started with a stimulus presentation phase of 1500 ms. The duration of the stimulus presentation was selected based on tests using pilot subjects and should ensure that enough trials can be recorded to get reliable estimates for the EEG measures used and the desired affective responses are evoked. During the stimulus presentation phase the subjects had to indicate if the current stimuli was a target (i.e. identical to the stimulus one step before in the 1-back condition and two steps before in the 2-back condition) or if the current stimulus was a distractor. Left and right control keys of a standard keyboard were used as inputs. The subjects were instructed to react as quickly as possible to ensure an effect in performance measures. Between stimuli there were interstimulus intervals (ISI) of 1500 ms duration including one to 500 ms of jitter at the end of the ISI to avoid periodic responses in the EEG data. During the ISI the same grey crosshair as in the baseline conditions was presented on the screen.

From psychological perspective, the N-back task can be divided into two phases. The first phase starts with the stimulus presentation and ends after the stimulus response. In this phase subjects had to perform a simple matching task by comparing the current stimulus with the stimulus stored in short term memory. In the second phase, ranging from the stimulus response to the next stimulus onset multiple executive functions are required for a correct response. Inhibition, shifting and updating represent the core executive functions as identified by Miyake and colleagues (Miyake et al., 2000). Inhibition is seen as the ability to suppress automatic responses that might arise during task processing. Shifting is referring to the concept of cognitive flexibility, making us capable to switch between different tasks. Updating refers to the continuous monitoring and changing of content in working memory. Content in short term memory needs to be updated via inhibition of the last stimuli in the stimuli list. Furthermore, the participants need to switch from a simple matching task to a working memory task and back. We therefore assume that the second phase is more relevant for the measurement of WML. Concerning the affective reaction we first assumed that it would be stronger right after stimulus onset, since the stimulus was fresh. However, pilot measurements revealed that subjects tend to focus first on the response toward the stimuli and only then direct their attention toward the affective content of the presented stimuli. We therefore decided to focus our analyses on the second phase.

Analysis

The window of analysis was 1400 ms wide. It started 1100 ms after stimulus onset to exclude post-motor responses in the EEG and ended 2500 ms after stimulus onset. The second time window also included 1000 ms of the ISI.

All spectra for the EEG analysis were computed using the Welch method (Welch, 1967) implemented in the EEGLAB toolbox (Delorme & Makeig, 2004). Theta bands were computed between 4 and 7 Hz and alpha bands were computed between 8 and 12 Hz. Power of the frontal (AFz, Fz) and parietal (CPz, Pz, POz) electrodes was averaged. Power for FAA computation was averaged across channels AF3, F3 and FC1 for the left hemisphere and across channels AF4, F4 and FC2 for the right hemisphere. We followed the approach from Allen, Coan, & Nazarian (Allen, Coan, & Nazarian, 2004) and computed FAA as difference score (see **Equation 1**).

Equation 1: Frontal Alpha Asymmetry Index (Allen et al., 2004)

$$\mathrm{FAA} = right\ alpha\ power - left\ alpha\ power$$

To evaluate the influence of affective valence and WML on the EEG, we performed repeated measures ANOVAs with two factors. The first factor was affective valence with three levels (positive, neutral and negative). The second factor was WML with two levels (1-back and 2-back). All analyses were conducted at the group level.

Results

Behavioral Performance Measures

Accuracies

Using accuracies as dependent variable we found a significant main effect for WML, $F(1, 23) = 32.70$, $p<.001$, $\eta2 = 0.59$. Higher WML resulted in decreased accuracies. Additionally, there was a significant main effect for affective valence, $F(2, 46) = 3.73$, $p < .035$, $\eta2 = 0.14$. A Post Hoc test using the Šidák correction revealed that the negative condition resulted in a decreased accuracy when compared to the positive condition ($p<.03$), but not when compared to the neutral condition (n.s.). Interestingly, there was a significant interaction between affective valence and WML, $F(2, 46) = 3.68$, $p<.035$, $\eta2 = 0.14$. The negative condition had a negative impact on accuracy, but only under high WML. Box plots showing the different conditions can be seen in **Figure 16** (left).

Reaction Times

There was a significant main effect for WML, $F(1, 23) = 60.53$, $p<.001$, $\eta2 = 0.73$. Higher load resulted in longer reaction times. Moreover, there was a significant main effect for affective valence, $F(2, 46) = 10.94$, $p<.001$, $\eta2 = 0.32$. The negative condition resulted in longer reaction times than the positive condition ($p<.01$) as well as the neutral ($p<.005$) condition. However, there was no significant interaction between affective valence and WML, $F(2, 46) = 1.02$, n.s. Median values for all conditions are shown via box plots in **Figure 16** (right).
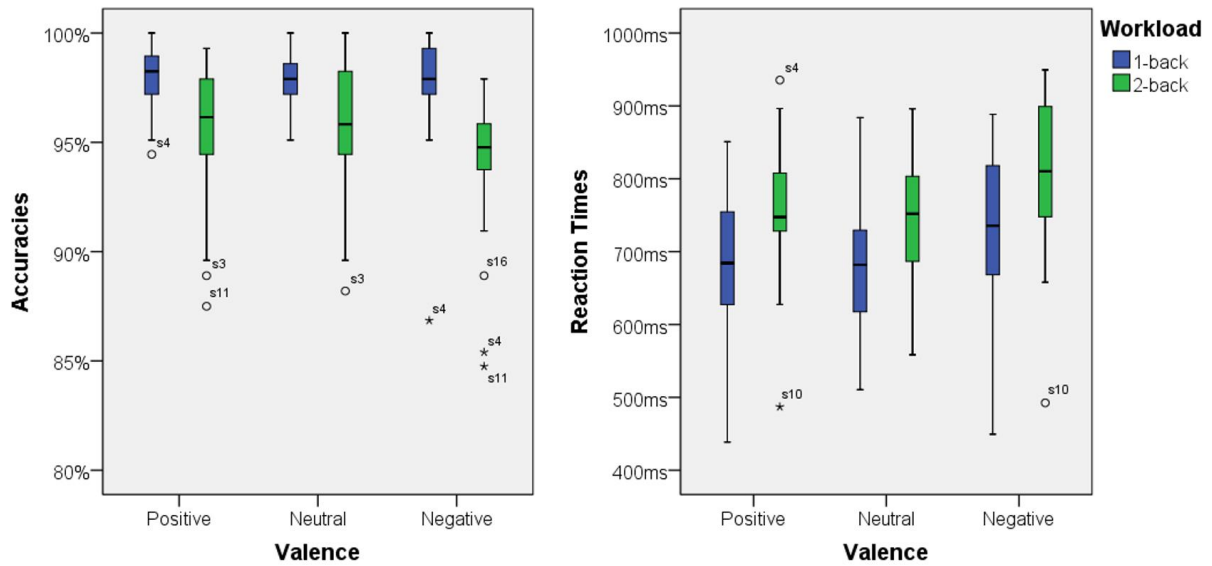
Figure 16. Box plots showing median values for accuracies and reaction times: Blue boxes show accuracies (left) and reaction times (right) for the 1-back conditions. Green boxes show accuracies (left) and reaction times (right) for the 2-back conditions. Median values are indicated by black horizontal lines within the boxes. Top and bottom borders of the boxes represent the middle 50% of the data. Whiskers represent the smallest and largest values not classified as outliers (between 1.5 and 3 times the height of the boxes) or extreme values (more than three times the height of the boxes). Circles indicate outliers and stars show extreme values.

EEG Working Memory Load Measures

Frontal Theta Activity

Using frontal theta activity as dependent variable there was a significant main effect for WML, $F(1, 23) = 9.00$, $p<.01$, $\eta2 = 0.28$. Frontal theta power was higher in the 2-back conditions. **Figure 17** (left) illustrates this effect. Interestingly, there was a significant main effect for affective valence, $F(2, 46) = 8.28$, $p<.001$, $\eta2 = 0.27$. A Post Hoc test using the Šidák correction revealed that frontal theta power was lower in the negative condition, when compared to the neutral condition ($p<.005$) as well as the positive condition ($p<.01$). **Figure 18** (left) shows a topographic plot displaying this frontal theta effect. However, there was no significant interaction between affective valence and WML, $F(2, 46) = 0.87$, n.s. Median values for all conditions can be seen in **Figure 19** (left).
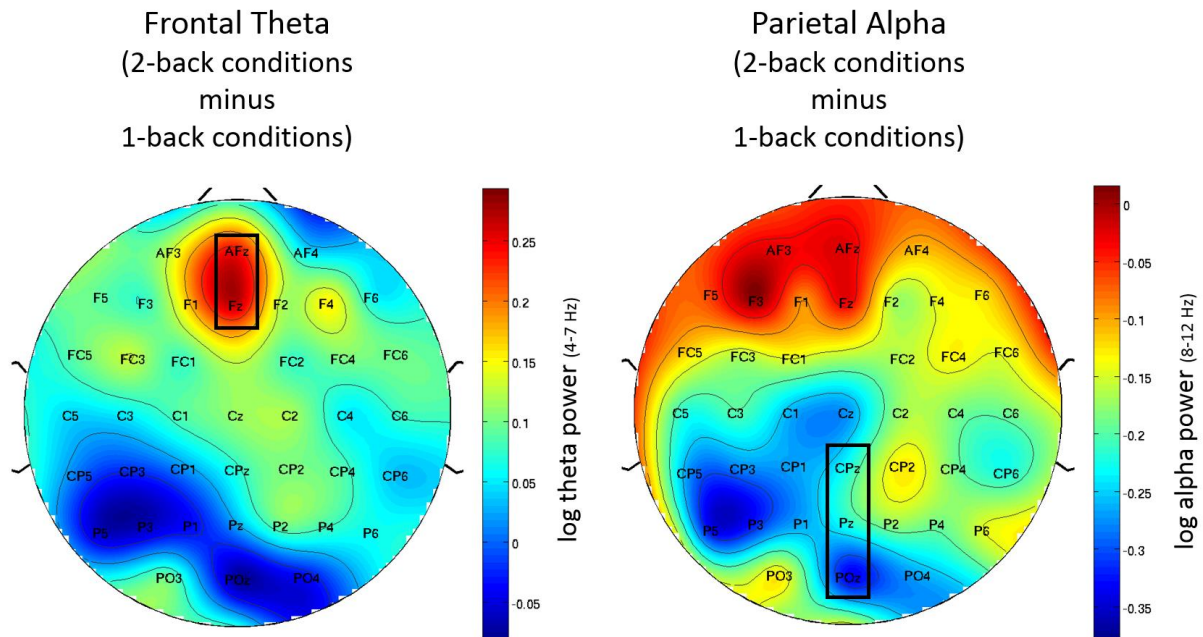
Frontal Theta
(2-back conditions
minus
1-back conditions)

Parietal Alpha
(2-back conditions
minus
1-back conditions)

Figure 17. Topographic plots showing difference in EEG workload measures between 2-back and 1-back conditions. Left: Frontal theta for the post-motor time window showing increased frontal theta activity for the 2-back conditions. Right: Parietal alpha for the Post motor time window showing decreased in parietal alpha power for the 2-back conditions. Nose is at the top. Values are averaged across all three affective conditions. Electrodes used for analysis are marked with black rectangles.

Parietal Alpha Activity

Using parietal alpha activity as dependent variable we found a significant main effect for WML, $F(1, 23) = 4.22$, $p=.05$, $\eta2 = 0.16$. Higher WML resulted in decreased parietal alpha power. This effect can be seen in **Figure 17** (right). Interestingly, there was also a significant main effect for the factor affective valence in the post-motor window, $F(2, 46) = 5.27$, $p<.01$, $\eta2 = 0.19$. The negative condition exhibited lower parietal alpha power than the neutral condition ($p<.03$), but not lower than the positive condition (n.s.). **Figure 18** (right) shows a topographic plot displaying this parietal alpha effect. There was no significant interaction between affective valence and WML, $F(2, 46) = 1.52$, n.s. **Figure 19** (middle) shows box plots for all conditions.
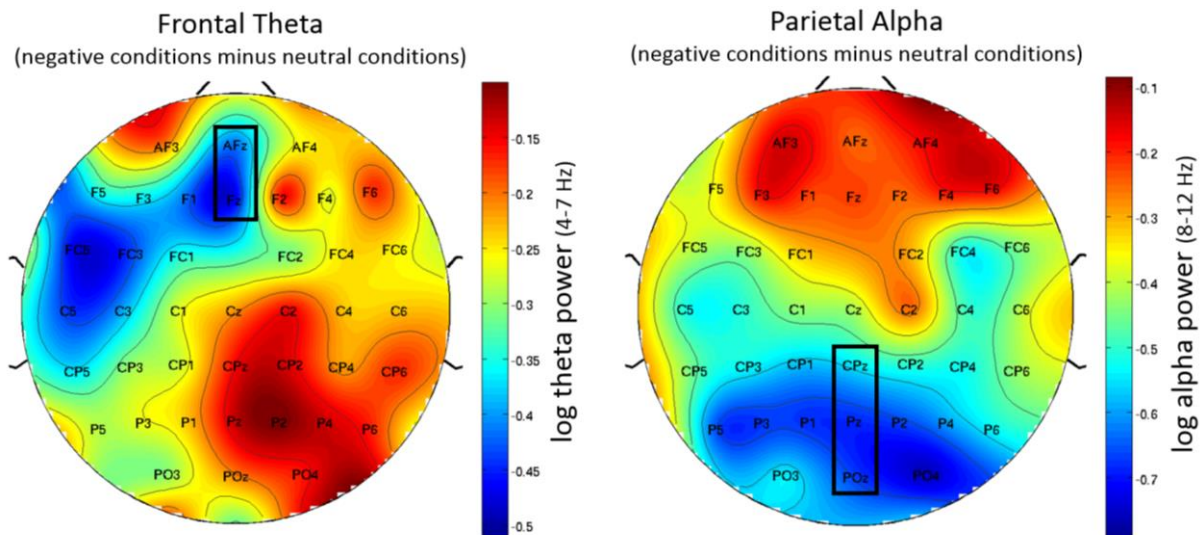
Figure 18. Topographic plots showing difference in EEG workload measures between negative and neutral conditions. Left: Frontal theta showing decreased frontal theta activity for the negative conditions when compared to the neutral conditions. Right: Parietal alpha showing widespread decrease in parietal alpha power for the negative conditions when compared to the neutral conditions. Nose is at the top. Values are averaged across both N-back levels. Electrodes used for analysis are marked with black rectangles.

EEG Affective Valence Measures – Frontal Alpha Asymmetry

Unexpectedly, there were no significant main effects with regard to FAA. Neither for affective valence, $F(2, 46) = 1.87$, n.s. nor for WML, $F(1, 23) = 0.96$, n.s. Finally, there was also no significant interaction between affective valence and working memory, $F(2, 46) = 0.08$, n.s. Median values for all conditions are summarized in **Figure 19** (right).
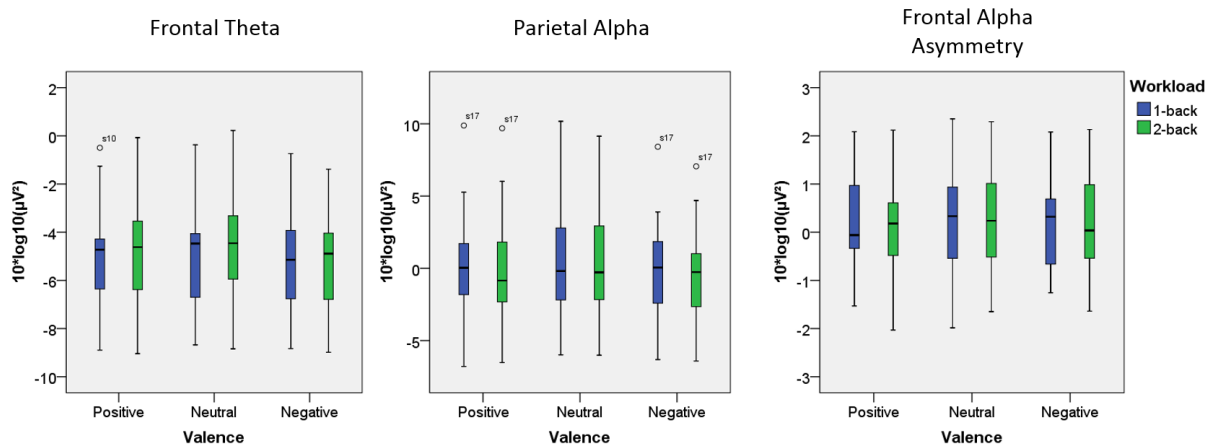
Figure 19. Box plots showing median values for frontal theta activity, parietal alpha activity and Frontal Alpha Asymmetry (FAA): Blue boxes show frontal theta activity (left), parietal alpha activity (middle) and FAA (right) for the 1-back conditions. Green boxes show frontal theta activity (left), parietal alpha activity (middle) and FAA (right) for the 2-back conditions. Median values are indicated by black horizontal lines within the boxes. Top and bottom borders of the boxes represent the middle 50% of the data. Whiskers represent the smallest and largest values not classified as outliers (between 1.5 and 3 times the height of the boxes) or extreme values (more than three times the height of the boxes). Circles indicate outliers. Please note the strong variability in the data due to large inter-individual differences.

Discussion

Using the emoback paradigm, we found that increased WML had a negative impact on task performance, as reflected in decreased accuracies and increased reaction times. These effects were also reflected in corresponding EEG measures. Increased WML was accompanied by increases in frontal theta activity as well as decreases in parietal alpha activity.

We found that negative affective valence had a negative impact on accuracies as well as reaction times. Interestingly, measures used for WML estimate also appeared sensitive to changes in affective valence. In contrast to that, FAA, a measure typically used to infer affective valence, did not show any effects in our paradigm, neither with regard to affective valence, nor with regard to increases in WML. In the following sections we will discuss these results in detail.

Performance

Our analyses revealed that increased WML induced via the emoback task did result in decreased performance. The 2-back conditions had reduced accuracy as well as increased reaction times. We also found that inducing negative affective valence had

detrimental effects on accuracies as well as reaction times. Similar results have been found in a study by Passarotti et al. (2011). The authors used face stimuli in an affective N-back task and found slower reaction times for angry faces. These negative effects of affective stimuli on cognitive processing can be explained within the theory of hot and cold cognition, concepts related to executive function (Zelazo & Muller, 2002). Hot cognition refers to a process where a person´s thinking is influenced by their affective state (Brand, 1986), while cold cognition is more based on rational thinking and critical analysis (Roiser & Sahakian, 2013). Hot cognition is able to overpower cold cognition in certain situations and has been shown to impair decision making (Huijbregts, Warren, De Sonneville, & Swaab-Barneveld, 2008). Processing of negative stimuli can divert cognitive resources from the primary task, and thereby lead to decreased performance. A study by Levens & Gotlib (2010) found that strongly valenced stimuli tend to stay longer active in working memory, which might interfere with different core executive functions necessary during the emoback task (Miyake et al., 2000). In the emoback task, subjects constantly needed to update content in working memory by replacing (inhibition) previous stimuli. Additionally, the subjects also needed to shift between these updating tasks and a rather simple identity-matching task. Negative stimuli could catch the attention, slowing down reorganization of stimuli in working memory and thereby impairing task performance.

Interestingly, we also found an interaction for accuracy between WML and affective valence. Negative affective valence resulted in decreased accuracies, but only during increased WML. These results are in line with findings from the study by Kopf et al. (2013) who found more errors for the difficult task during the negative condition using an affective word N-back task. These results seem to indicate that both, cognitive and affective processing, compete for limited cognitive resources and this dual strain on mental resources leads to decreased performance. Similar results have also been found in the study by MacNamara et al. (2011). The authors used emotional pictures as distractors in an emoback task and found that the emotional content of negative valenced stimuli increased the negative impact of WML on performance. These findings could be explained using the capacity model from Ellis & Ashbrook (1989). Their model assumes that there is a limited pool of attentional resources that can be used to complete a certain task. Affective states can influence the allocation of available attentional resources toward the task, thereby potentially impairing task

performance. However, based on performance measures alone, we can only vaguely assume how negative valence impaired task performance.

EEG Working Memory Load Measures

Frontal Theta Activity

Increased WML was reflected in EEG measures as well. Frontal theta activity was increased during high WML. Interestingly, the negative affective valence condition exhibited decreased frontal theta after the motor response. Additional WML due to the processing of negative affective stimuli should have been reflected in increased frontal theta activity. Decreased frontal theta activity seems to imply decreased processing in frontal executive areas.

Parietal Alpha Activity

Our analyses revealed that parietal alpha activity was decreased under high WML (2-back) condition. We also found that the parietal alpha activity was reduced for the negative condition. An fMRI study by Rämä et al. (2001) used different affective voices and found similar results. The authors discovered that the parietal cortex is bilaterally involved in active maintenance of emotional content. Affective stimuli are more salient and usually carry more relevant information (Carretié, 2014). We therefore assume that negative valenced stimuli result in stronger processing in the storage areas of the parietal cortex, which seems reflected in increased parietal activity (i.e. reduced alpha power).

EEG Affective Valence Measures – Frontal Alpha Asymmetry

Unexpectedly, we did not find any effect using the FAA measure. We used the FAA measure during the induction of WML, while most other studies used brain laterality responses during rest or passive viewing (Ahern & Schwartz, 1985; Davidson & Wheeler, 1992; Huang et al., 2012; Lin, Wang, Wu, Jeng, & Chen, 2009; Ramirez & Vamvakousis, 2012). Use of the FAA in combination with other tasks is further complicated because the FAA is known to be also influenced by other factors like seating position and WML (Briesemeister, Tamm, Heine, & Jacobs, 2013). A study by Baldwin et al. 2012 (Baldwin & Penaranda, 2012) found that increased task difficulty resulted in increased left frontal activity. It appears that the negative condition increased task difficulty, since it did impair task performance. Negative affective valence in passive conditions usually results in increased right frontal activity. It is therefore conceivable that both, affective and workload related, processes acted on

the same FAA measure, but in opposing directions. This could mean that both effects canceled each other out and thereby masked any potential effects.

Limitations and Outlook

We did not find any effects using the FAA measure. We assume that was due to the WML elicited through our paradigm. It would be interesting to test this assumption by developing a study design that contrasts an active condition with a passive viewing condition.

Furthermore, future studies should explore how well these findings generalize to male individuals as well as other stimulus modalities.

Conclusion

We demonstrated that negative valenced stimuli can have an effect on performance measures. Additionally, when using established EEG measures we found that increased WML can be detected in the EEG, even when affective valence is induced at the same time. Although FAA did not prove useful for identification of emotional states, our analyses revealed interactions between affective valence and WML. Future studies should further investigate the context sensitivity and applicability of EEG measures in various contexts to identify the ramifications in which such measures can be used to identify different states and processes.

6.1.1 Additional Chapter: Subjective Measures Results

In this chapter I will present the results from the analysis of the subjective ratings recorded in the emoback study.

Workload Ratings

Subjective experience of WML was measured using one (modified) item taken from the NASA task load index (Hart & Staveland, 1988). The item asked participants how

cognitively demanding they experienced the last experimental run. The rating scale ranged from zero (absolutely no mental demand) to one-hundred (highest possible mental demand).

As expected, using workload ratings as dependent variable, we found a significant main effect for WML, $F(1, 23) = 22.7$, $p<.001$, $\eta^2 = 0.50$. Higher WML resulted in the increased subjective experience of WML. Additionally, there was a significant main effect for affective valence, $F(2, 46) = 11.6$, $p < .001$, $\eta^2 = 0.34$. A Post Hoc test using the Šidák correction revealed that the negative valence condition resulted in an increase of subjective WML when compared to the positive valence condition ($p<.001$) as well as the neutral valence condition ($p<.02$). There was no significant interaction between affective valence and WML, $F(2, 46) = 0.8$, n.s. Box plots showing the different conditions can be seen in **Figure 21** (left).

Affective Valence Ratings

Emotional experiences within the PAD are commonly judged with the help of rating scales. The most widely used is a visual analog scale called the Self-Assessment Manikin (SAM; Bradley & Lang, 1994). It enables fast and reliable judgements about the current emotional state. Due to the use of simple depictions it can also be utilized with children, which could be an additional benefit in the context of the development of a neurotutor. **Figure 20** shows a depiction of the three SAM scales for valence, arousal and dominance.
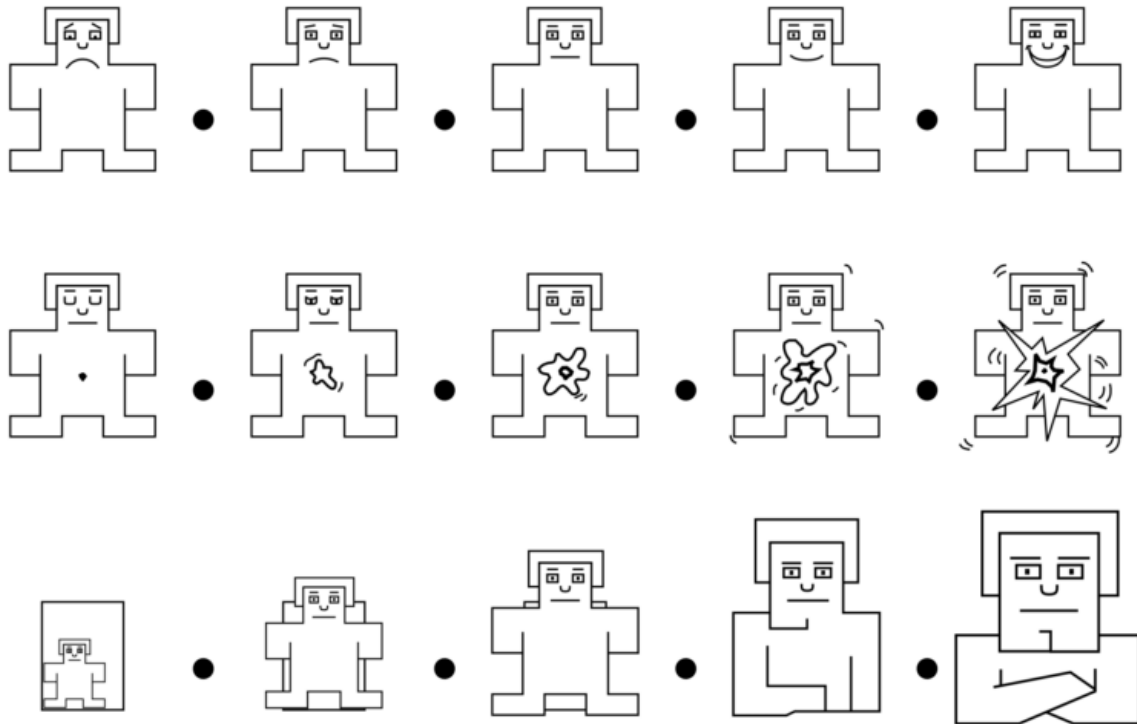
Figure 20. The Self-Assessment Manikin. Top: The valence dimension ranging from extreme displeasure (left) to extreme happiness (right). Middle: The arousal dimension ranging from extreme relaxation (or sleep) to frenzied excitement (right). Bottom: The dominance or control dimension ranging from low dominance (left) to high dominance (right; Bradley & Lang, 1994).

Arousal ratings were not analyzed, since the main focus of this thesis are responses related to the valence or dominance dimension. Additionally, participants reported strong difficulties with regard to the dominance dimension and therefore mostly chose the middle category. Accordingly, I only analyzed ratings from the valence dimension.

As anticipated, using the affective valence ratings as dependent variable, there was a significant main effect for affective valence, $F(1.4, 33) = 55.7$, $p<.001$, $\eta^2 = 0.71$. Decreased affective valence due to the emotion induction resulted in decreased subjective valence ratings. The neutral condition resulted in more positive valenced ratings than the negative condition ($p>.001$) and more negative valenced ratings than the positive condition ($p>.01$). However, there was neither a significant main effect for WML, $F(1, 23) = 0.1$, n.s., nor was there a significant interaction between affective

valence and WML, F(2, 46) = 1.0, n.s. Median values for all conditions are shown via box plots in **Figure 21** (right).
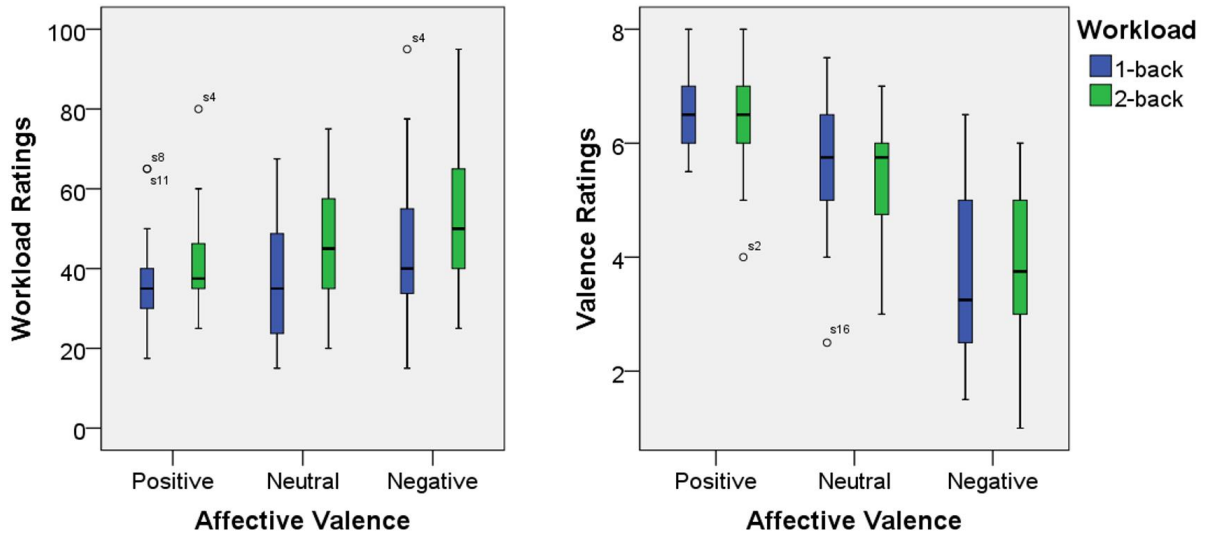
Figure 21. Box plots showing median values for workload ratings (left) and valence ratings (right): Blue boxes show workload ratings (left) and valence ratings (right) for the 1-back conditions. Green boxes show workload ratings (left) and valence ratings (right) for the 2-back conditions. Median values are indicated by black horizontal lines within the boxes. Top and bottom borders of the boxes represent the middle 50% of the data. Whiskers represent the smallest and largest values not classified as outliers (between 1.5 and 3 times the height of the boxes) or extreme values (more than three times the height of the boxes). Circles indicate outliers.

6.2 Context Sensitivity of EEG-Based Workload Classification under different Affective Valence (Emoback Study 2)

Sebastian Grissmann[1]*, Martin Spüler[2], Josef Faller[3], Tanja Krumpe[2], Thorsten O. Zander[1,4], Augustin Kelava[1,5], Christian Scharinger[6] & Peter Gerjets[1,6]

[1] LEAD Graduate School, University of Tübingen, Tübingen, Germany

[2] Wilhelm-Schickard-Institute for Computer Science, University of Tübingen, Tübingen, Germany

[3] Laboratory for Intelligent Imaging and Neural Computing, Columbia University, New York, USA

[4] Team PhyPA, Biological Psychology and Neuroergonomics, Berlin Institute of Technology, Berlin, Germany

[5] Hector Research Institute of Education Science and Psychology, University of Tübingen, Tübingen, Germany

[6] Leibniz-Institut für Wissensmedien, Tübingen, Germany

Abstract

State of the art brain computer interfaces (BCIs) largely focus on detecting single, specific, often experimentally induced or manipulated aspects of the user state. We argue, that in a less controlled, more naturalistic environment, a larger variety of mental processes may be active and possibly interacting. When moving BCI applications from the lab to real-life applications, these additional unaccounted mental processes could interfere with user state decoding, thus decreasing system efficacy and decreasing real-world applicability. Here, we assess the impact of affective valence on classification of WML, by re-analyzing a dataset that used an affective N-back task with picture stimuli. Our analyses showed that classification of WML under affective valence can lead to good classification accuracies (> 70%), which can be further improved via integration over time. However, positive as well as negative affective valence resulted in decreased classification accuracies. Furthermore, classifiers failed to generalize across affective contexts, highlighting the need for frameworks that can account for different contexts or new, context-independent, EEG features.

Introduction

Brain-based Adaptive Systems

The electroencephalogram (EEG) represents an interesting method for the evaluation of human machine interaction (Frey, Mühl, Lotte, & Hachet, 2014) by enabling the design of adaptive, brain-based human-machine interaction systems. Such adaptive systems could allow us to tailor future workspaces to the individual needs of users. Passive brain-computer interfaces (BCIs) represent a good example for such adaptive systems. Passive BCIs utilize mental state monitoring to establish an implicit control loop and therefore do not rely on awareness of the user about this communication. This implicit control loop could use additional context information about the user to potentially improve a given human-machine interaction. Passive BCIs have been previously applied to adaptive learning environments (Walter et al., 2013) or even detection of covert user states like bluffing in a game context (Reissland & Zander, 2009).

Challenges when moving out of the Lab

Most current passive BCIs are used in lab environments and focus on narrow user states like WML (Spüler et al., 2016) or affective valence (Smedt & Menschaert, 2012). Therefore, such systems might not be able to provide stable performance outside the lab, since real world environments tend to evoke a combination of multiple basic mental processes (Gerjets et al., 2014). Furthermore, many measures used to infer mental states from the ongoing EEG are known to have many-to-many relations, meaning that several physiological variables are associated with multiple psychological elements (Fairclough, 2009). Previous work suggests that contextual changes, as expected to arise in natural environments, could be related to spontaneously occurring decrements in passive BCI performance (Myrden & Chau, 2015; Zander & Jatzev, 2012). Most available classification algorithms are based on stable signal properties, which are derived from a single calibration session. Such classification approaches therefore require the user´s brain to adapt to the new behavior of the system, which can be very challenging, since we are usually not aware how to directly modulate EEG-features used for BCI classification without extensive training (Birbaumer, 2006). Therefore such classification approaches might have difficulties to track shifts in the feature space due to contextual changes of the current user state. Newly developed approaches try to mitigate this issue with methods from machine learning like adaptive

classification schemes (Faller, Vidaurre, Solis Escalante, Neuper, & Scherer, 2012; Shenoy, Krauledat, Blankertz, Rao, & Müller, 2006; Spüler, Rosenstiel, & Bogdan, 2012). Such approaches usually work well in controlled lab environments, but to make them feasible outside the lab, we want to investigate if and how certain EEG features used for mental state detection are influenced by specific contexts. To identify such context-sensitivity in EEG features used for passive BCI classification we investigated a combination of two important mental states used in passive BCIs, WML and affective valence.

Working Memory Load

Automatic detection of WML has already been studied extensively for brain-based adaptive systems (Berka et al., 2007; Kohlmorgen et al., 2007; Walter et al., 2013). Working memory is usually defined as the interaction of short term memory structures that handle different representational codes with controlled attentional processes (Baddeley, 2003, 2012). Content in short term memory is thought to be maintained via parietal brain areas like the intraparietal sulcus, while the central attentional control system is assumed to be mainly located in frontal regions of the brain like the dorsolateral prefrontal cortex (Klingberg, 2009; Scharinger et al., 2015). Accordingly, increases in mental workload commonly result in decreased parietal alpha activity as well as increased frontal theta activity (Gevins et al., 1995; Klimesch, 1999; Smith & Gevins, 2005). However, previous research has shown that other mental states can also have an effect on measures used for workload detection. For instance, Roy et al. (2013) induced fatigue and found that with increasing time on task the discriminability of WML was decreased.

Affective Valence

The circumplex model (Russell, 1980) is a popular, theoretical framework that describes affective states with the dimensions valence and arousal. On one hand, the arousal dimension reflects the responsiveness of the organism and ranges from sleep to frenzied excitement. On the other hand, the valence dimension reflects the pleasantness of a situation and ranges from sadness to happiness.

The Emoback Paradigm

To experimentally modulate WML and affective valence separately we decided to use the Emoback paradigm. The emoback paradigm combines an N-back task with an

affect induction. The N-back task is a continuous performance task where subjects are presented with a series of stimuli and have to indicate whether the current stimulus is identical to the stimulus presented N-steps before, or not (Kirchner, 1958; Scharinger et al., 2015). Hence, the load factor 'N' allows to adjust the difficulty of the task, thereby manipulating WML.

There are various ways to induce emotions in lab settings (Gerrards-Hesse et al., 1994). One efficient method is the use of affective picture stimuli. One very popular picture database is the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 1997) which comprises a large set of standardized and emotionally evocative color photographs. All stimuli of the database have been rated along the dimensions of valence and arousal using the self-assessment manikin (SAM; Bradley & Lang, 1994).

Other paradigms, like the affective flanker (see Alguacil, Tudela, & Ruz, 2013) have been used in the past to investigate potential influences of affect on EEG features. However, the affective flanker task only requires perceptual inhibition and does not involve a memory component. Furthermore, the emoback paradigm has, as far as we know, only been used three times in combination with the EEG. First, a study by MacNamara, Ferri, & Hajcak (2011) used the emoback task with affective pictures as distractors and found that the late positive potential (LPP) was modulated not only by affective responses toward the emotional pictures, but also with regard to changes in WML. However, the LPP is an EEG feature in the time domain and responds mainly to affective stimuli related to the arousal dimension (Schupp et al., 2000). Additionally, the study by MacNamara, Ferri, & Hajcak (2011) only used affective distractors, while we want to investigate a task which intrinsically activates cognitive and affective processes.

Second, a study by Kopf et al. (2013) used a N-back task with emotional words and recorded data from the EEG and the fNIRS. They found more errors during the negative condition, especially for the difficult task. An ERP-analysis also revealed that the LPP could be influenced by the difficulty in a working memory task and that this influence is further modulated by affective valence. However, we assume that emotional words evoke rather weak responses and cannot convey the complexity of real-life emotions.

Third, Grissmann et al. (under review) used an emoback task with affective picture stimuli and found that performance decreased with increasing workload as well as negative valence. These findings were reflected in changes in frontal theta and parietal alpha power. Interestingly, these EEG measures were also influenced by negative affective valence. However, the authors did not find any effects for EEG measures typically used for affective valence detection.

These studies show that EEG features can be context-sensitive and prone to influences not related to the primary mental state of interest. To extend these findings to measures used for detection of WML we decided to re-analyze the data from Grissmann et al. (under review) and investigate the automatic detection of WML under affective valence.

Research Questions

To be more precise, we investigated if classification of WML under affective valence can lead to useful classification accuracies. Additionally, we investigated if accuracies can be improved via data integration over time, which has been shown to work in a previous study (Brouwer et al., 2012). Finally, we also investigated the context-specificity of the resulting workload classifiers.

Methods

Sample

We re-analyzed data from Grissmann et al. (under review). They collected data from 27 female, college-aged (mean: 23.0 years; range: 19 to 32 years), right handed volunteers and rejected three data sets due to insufficient data quality. Subjects with blood phobia were excluded to avoid extreme responses toward the experimental stimuli. All subjects provided written informed consent and were paid 20 € for their participation. The study was approved by the ethics committee of the Knowledge Media Research Center Tübingen.

Recording of Physiological Data

Sixty channels of EEG were recorded using an ActiCHamp amplifier and active Ag/Cl-electrodes (Brainproducts GmbH, Gilching, Germany). Electrodes were placed according to the extended 10-20 system. The electrooculogram (EOG) was recorded

with four EEG electrodes located at the left and right canthi as well as above (channel Fp1) and below the left eye. All channels were referenced to channel FCz and impedances were kept below 10 kΩ. The data were sampled at 1 kHz. For later processing the data was downsampled to 500 Hz. During the recordings subjects were instructed to sit in a relaxed posture to avoid artifact contamination of the data.

Preprocessing

To automatically reject time windows that are contaminated by artifacts, a time window was removed if channel power at more than six channels exceeded five times the standard deviation. After computing independent components using the CUDAICA implementation (Raimondo, Kamienkowski, Sigman, & Fernandez Slezak, 2012) of the Infomax independent component analysis (Bell & Sejnowski, 1995), eye movement artifacts were removed (reduced) using the ADJUST approach (Mognon, Jovicich, Bruzzone, & Buiatti, 2011) which is implemented as plug-in in the EEGLAB toolbox (Delorme et al., 2011). After visual inspection of the resulting data, we decided to exclude channels located on the outer circumference of the electrode cap (channels Fp1, Fp2, AF7, AF8, F7, F8, FT7, FT8, T7, T8, TP7, TP8, P7, P8, PO7, PO8, O1, Oz and O2) from later analysis, since those channels showed artifactual activity. Finally the EEG signal was bandpass filtered between 1 and 45 Hz.

Stimuli

For the positive, neutral and negative valence conditions we selected 96 stimuli, each taken from the IAPS database (Lang et al., 1997). Valence and arousal ratings are usually confounded, meaning that stimuli with strong (positive or negative) valence ratings usually also have high arousal ratings (Olofsson, Nordin, Sequeira, & Polich, 2008). To improve discriminability between affective conditions we made sure that stimuli for the positive condition had the highest valence ratings while stimuli for the negative conditions had the lowest valence ratings using valence and arousal ratings provided with the database. Furthermore, stimuli for the neutral condition were selected based on the lowest arousal ratings. All selected stimuli had a quadratic shape and were centrally presented on a standard 23 inch display. To avoid unnecessary eye-movements the size of the stimuli was scaled to fill 60 % of the height of the display.

Study Design and Experimental Run Structure

All affective valence conditions (positive, neutral and negative) were presented in groups of four blocks with identical affective valence. Load levels (1-back and 2-back) of the emoback task were alternated run-wise. We avoided the use of a zero-back condition, because this would require to repeatedly use the same affective target stimuli. Such repetitions of the same stimuli might have resulted in affective habituation with regard to the target stimuli, thereby diminishing the intensity of the affect induction (Leventhal, Martin, Seals, Tapia, & Rehm, 2007). We also avoided a 3-back condition because some studies showed that this load level can lead to task disengagement (Ayaz, Izzetoglu, Bunce, Heiman-Patterson, & Onaral, 2007). Target response hand as well as starting load level of the emoback task were balanced across the subject sample (see **Figure 22**).
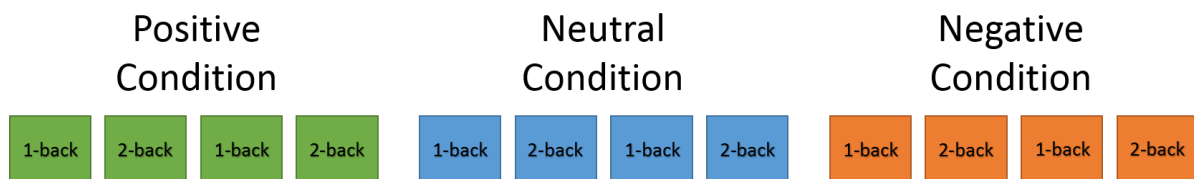
Figure 22. Illustration of emoback study design: The three affective valence conditions were presented in a block design and permutated across the whole sample. Each affective condition consisted of four workload runs, resulting in 12 runs for each subject. Workload runs were presented in an alternating fashion and balanced across the whole sample.

All participants performed 12 runs in total. Each run started with a 10 second baseline where subjects were instructed to relax and visually fixate a centrally presented light grey fixation cross (see **Figure 23**). After the baseline the first trial started with a stimulus presentation. There were 72 trials in one run, consisting of 24 target stimuli and 48 distractors. The 24 target stimuli were randomly selected for each subject from 96 unique stimuli (four blocks with 24 target stimuli per affective condition). Targets and distractors were sometimes interleaved in the 2-back condition, meaning that two target stimuli could appear right after each other. Here is an example: Distractor (e.g. table), distractor (e.g. chair), target (table) and target (chair). After the last stimulus (i.e. trial), another baseline phase, also with a duration of 10 seconds, was recorded. Between runs there were short brakes between one and three minutes.

1-back condition
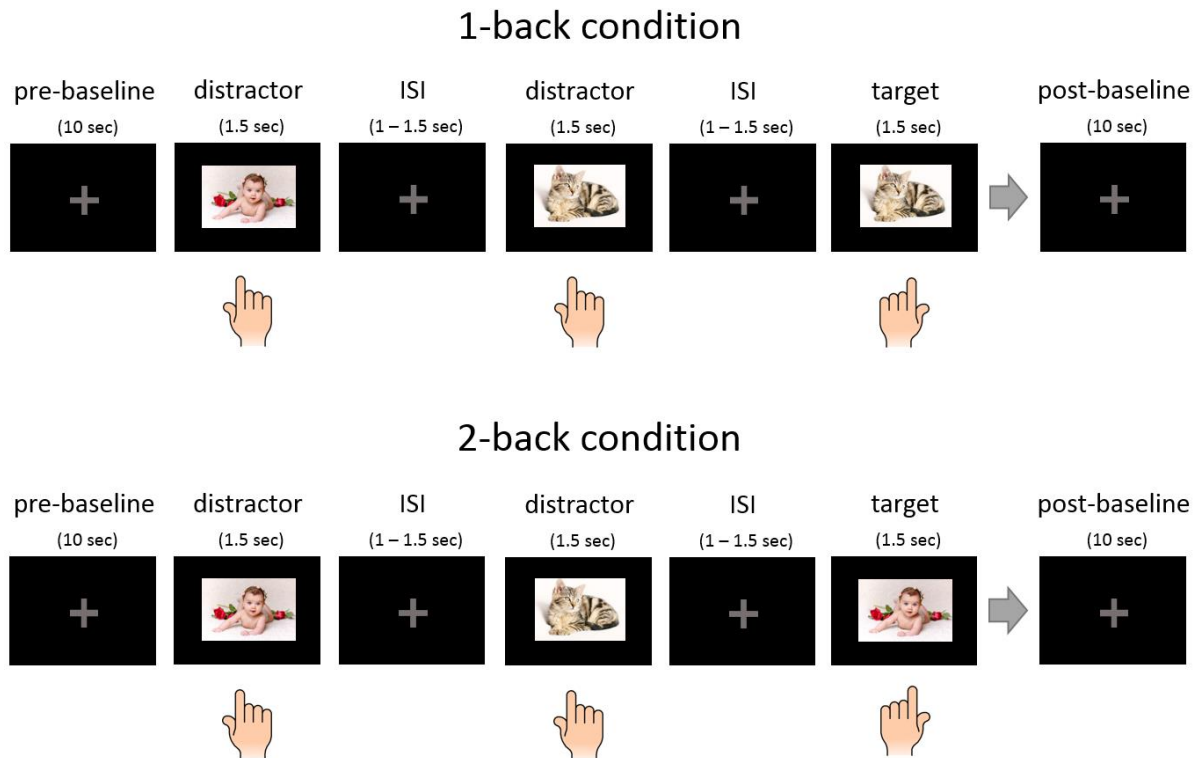
2-back condition

Figure 23. Schematic representation of emoback run structure. Top: 1-back condition. Bottom: 2-back condition. Hand symbols indicate key press. Left hand press was used for target stimuli and right hand press for distractors in this example. Pictures in this figure are taken from https://pixabay.com/ and only serve as illustration.

Trial Structure

Each trial started with a stimulus presentation phase of 1500 ms. The duration of the stimulus presentation was selected based on tests using pilot subjects and should ensure that enough trials can be recorded to get reliable estimates for the EEG measures used and the desired affective responses are evoked. During the stimulus presentation phase the subjects had to indicate if the current stimuli was a target (i.e. identical to the stimulus one step before in the 1-back condition or two steps before in the 2-back condition) or if the current stimulus was a distractor. Left and right control keys of a standard keyboard were used as inputs. The subjects were instructed to react as quickly as possible to increase chances for a potential effect in performance measures. Between stimuli there were interstimulus intervals (ISI) of 1500 ms duration including one to 500 ms of jitter at the end of the ISI to avoid periodic responses in the

EEG data. During the ISI the same grey crosshair as in the baseline conditions was presented on the screen.

Offline Analysis

The time window used for analysis was 2400 ms wide. It started 100 ms after stimulus onset to minimize exogenous responses and ended 2500 ms after stimulus onset. The time window also included 1000 ms of the ISI. All spectra for the EEG analysis were computed using an autoregressive model of order 32 with the Burg method (Priestly, 1982). Band power between 1 15 Hz in 1 Hz bins was used in the offline analysis. To determine which features might allow a good discrimination between different WML conditions, r2-values (Sheikh, McFarland, Sarnacki, & Wolpaw, 2003) were calculated between all relevant combinations of two WML conditions for each frequency bin at each electrode. Theta bands were computed between 4 and 7 Hz and alpha bands were computed between 8 and 12 Hz.

For classification a Support Vector Machine (SVM; Vapnik, 1999) implemented in the LibSVM Toolbox (Chang & Lin, 2011) with a linear kernel and the regularization parameter set to 1 was used. To estimate accuracies a 10x10-fold cross-validation in which the data was permuted and partitioned into 10 blocks of equal size was used. In each of the 10 folds, we used nine blocks for training the classifier and tested on the one remaining block. Each block was used for testing once. This procedure was repeated 10 times, and the accuracy was averaged over all folds. For the majority voting we used the probability estimate from the resulting SVM classifier and later averaged across multiple trials to get an improved workload estimate.

To evaluate the influence of affective valence on WML classification in the EEG, we performed a repeated measures ANOVA with two factors and classification accuracies as dependent variable. The first factor was affective valence with three levels (positive, neutral and negative). The second factor was number of trials with five levels, representing single trial classification as well as majority voting across multiple (up to five) trials. This factor was used to investigate potential improvements on classification accuracies when using larger time segments for classification. Especially passive BCI approaches could benefit from such an approach, since stable state prediction is usually more important than fast reaction times of the system. All

classifications were conducted at the single subject level and results were later averaged across the whole sample.

To further investigate how workload classification is influenced by a specific context (in this case affective valence) we used a cross-valence classification approach, where we trained the SVM classifier on WML within one affective context and later tested the classifier under a different affective context. More specifically, we trained a workload classifier under positive affective valence and tested the resulting workload classifier under negative affective valence, and vice versa. Furthermore, we wanted to investigate if the decision of the workload classifier is biased by the underlying affective condition. We therefore trained a workload classifier under low affective valence (neutral conditions) and later tested it on two affective valence conditions (positive 1-back Vs negative 1-back).

Results

All mean WML classification accuracies were above 70 %, even with single trial classification and did go as high as 87.7 % via data integration across five trials (all median values are shown via box plots in **Figure 24**).
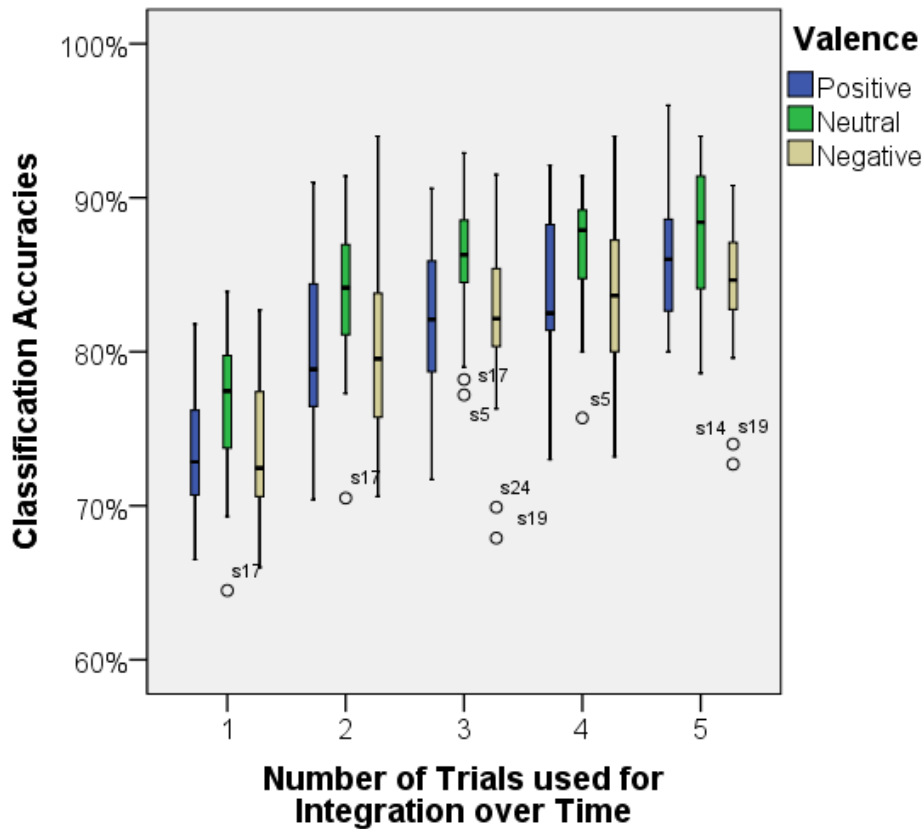
Figure 24. Box plots showing median values for accuracies: Blue boxes show accuracies for WML under positive affective valence. Green boxes show accuracies for WML under neutral affective valence. Beige boxes show accuracies for WML under negative affective valence. Median values are indicated by black horizontal lines within the boxes. Top and bottom borders of the boxes represent the middle 50% of the data. Whiskers represent the smallest and largest values not classified as outliers (between 1.5 and 3 times the height of the boxes) or extreme values (more than three times the height of the boxes). Circles indicate outliers and stars show extreme values.

For the discrimination of workload conditions under positive affective valence, parietal alpha features were most important, however frontal alpha features were involved as well. Discrimination of workload conditions under neutral affective valence was mainly based on parietal alpha features (see **Figure 26**), while discriminability for the workload conditions under negative affective valence depended on frontal theta features as well as parietal alpha features (see **Figure 27**). Interestingly, we found that r^2 plots for the workload conditions under positive as well as neutral affective valence showed some parietal theta and frontal alpha activity (see **Figure 25** and **Figure 26**).
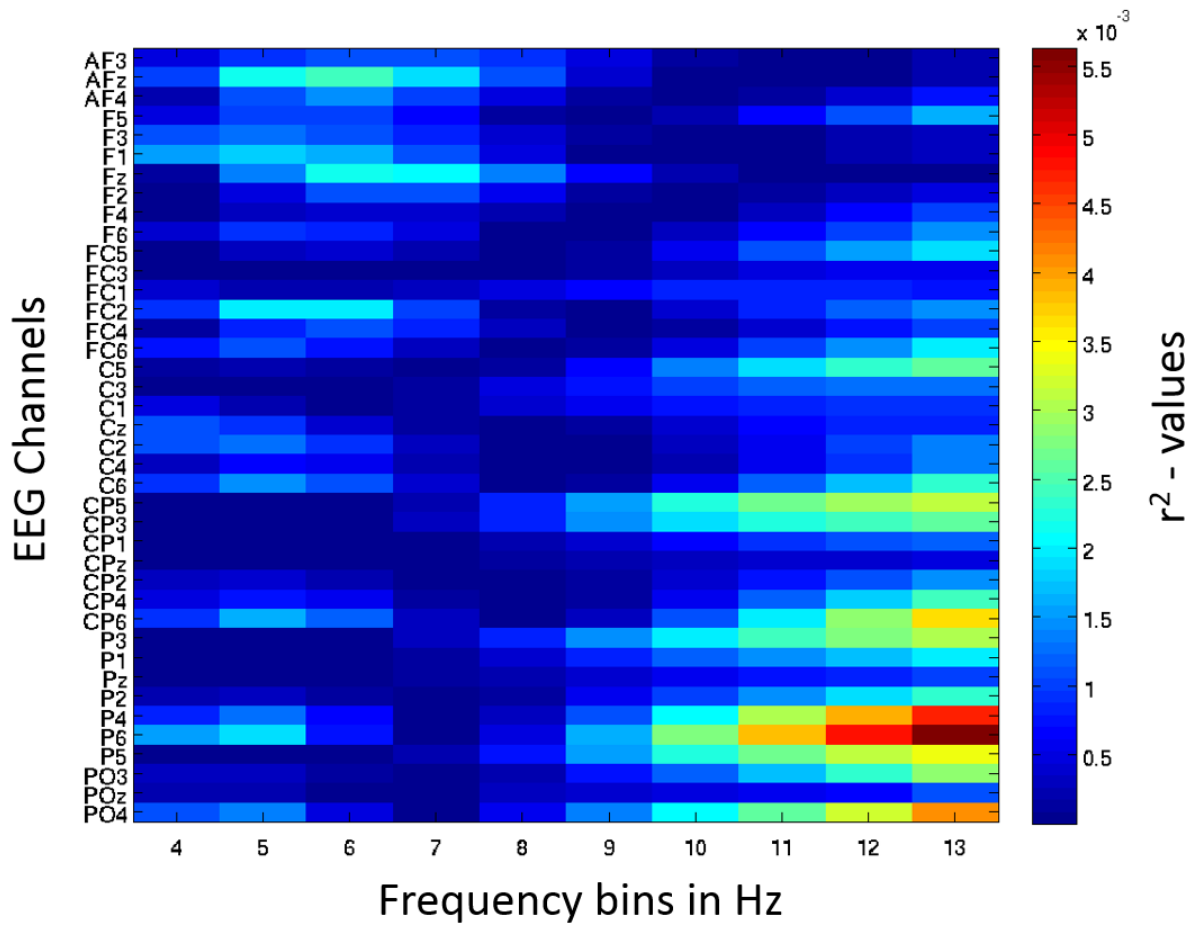
Figure 25. R² plot for the 1-back and 2-back conditions under positive affective valence.
Rows show individual EEG channels and Columns show individual frequency bins. R² values
are color coded (see colorbar on the right side of figure for scaling).
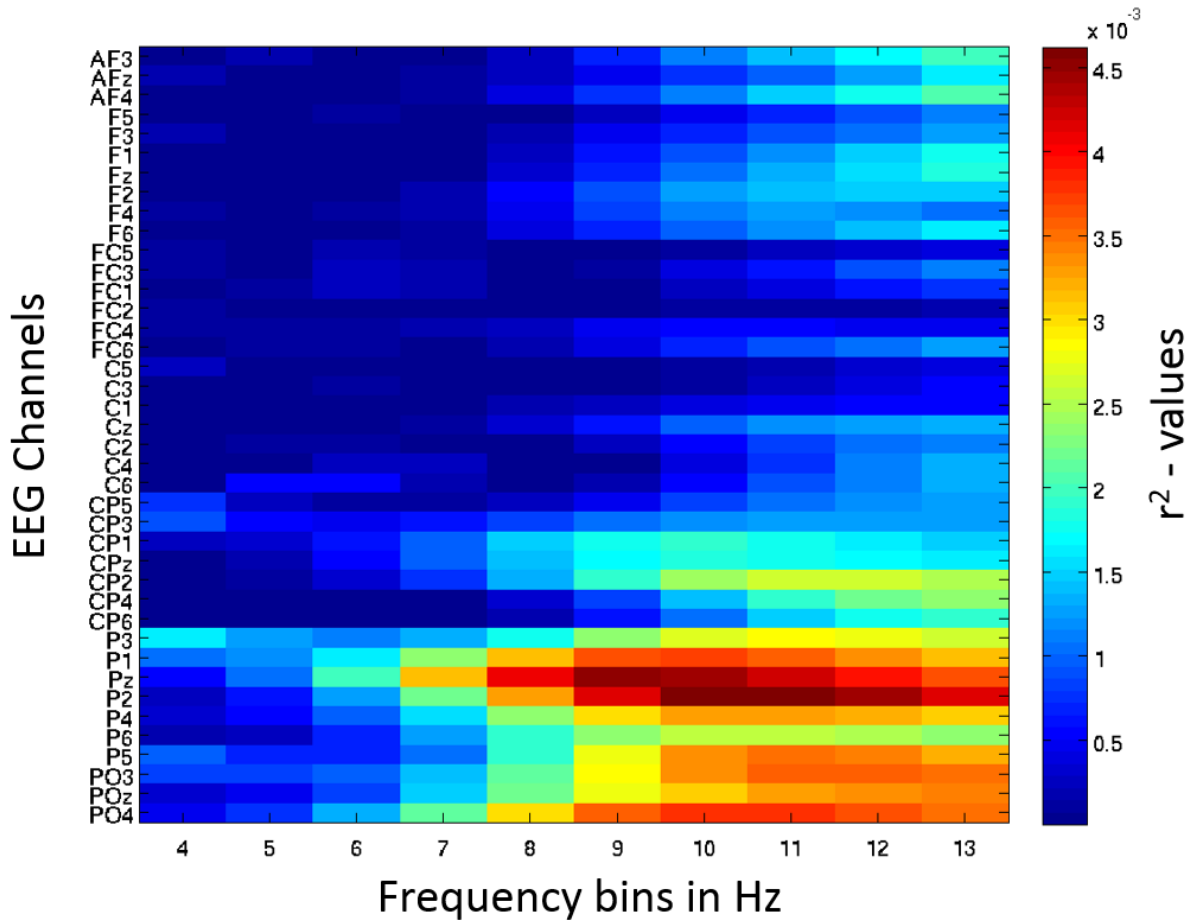
Figure 26. R² plot for the 1-back and 2-back conditions under neutral affective valence. Rows show individual EEG channels and Columns show individual frequency bins. R² values are color coded (see colorbar on the right side of figure for scaling).

Negative 1-back Vs Negative 2-back
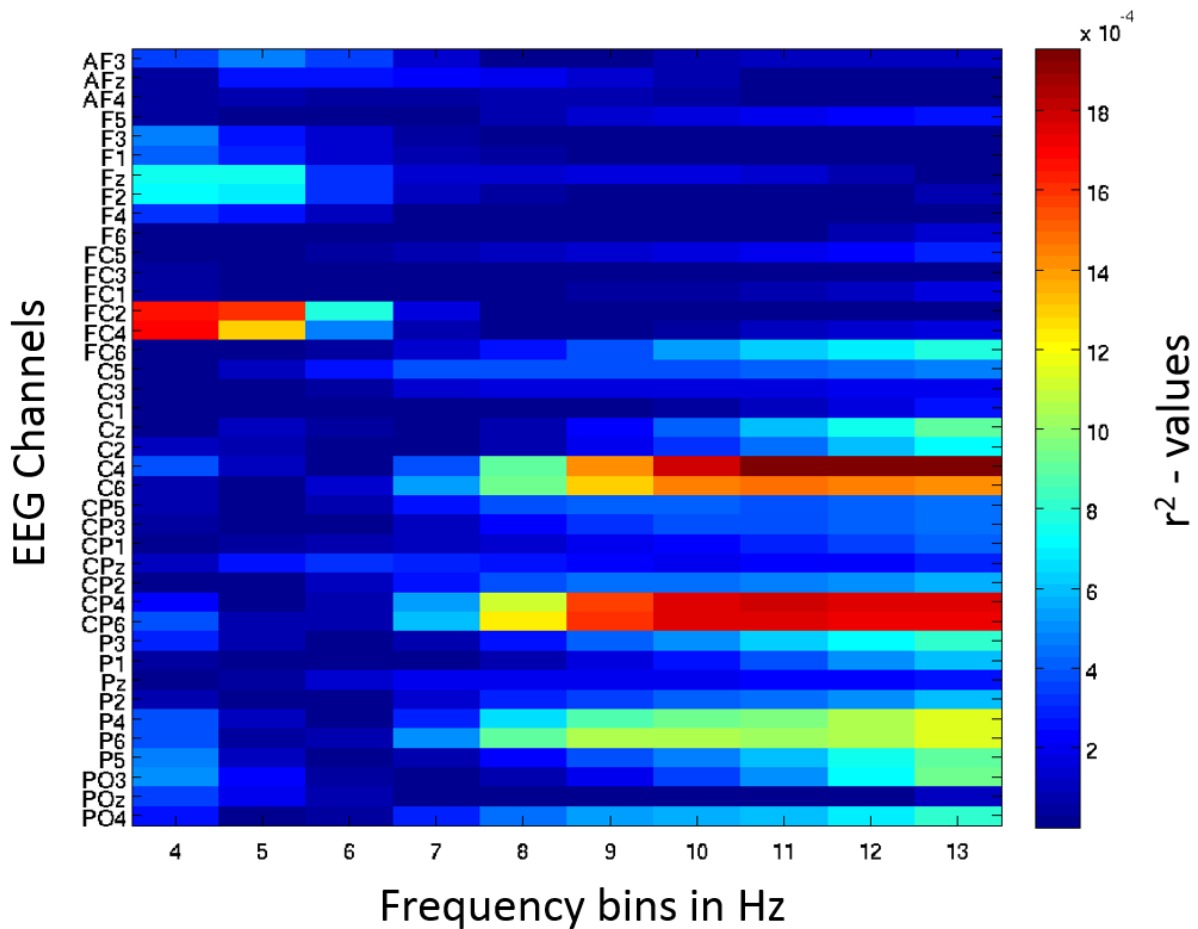
Figure 27. R² plot for the 1-back and 2-back conditions under negative affective valence. Rows show individual EEG channels and Columns show individual frequency bins. R² values are color coded (see colorbar on the right side of figure for scaling).

Interestingly, using classification accuracies as dependent variable there was a significant main effect for affective valence, $F(2, 46) = 5.0$, $p<.015$, $\eta^2 = 0.18$. A Post Hoc test using the Šidák correction revealed that the neutral condition resulted in higher accuracies than the negative condition ($p<.05$) as well as the positive condition ($p<.05$). As expected, we found a significant main effect for number of trials, $F(2.8, 64.6) = 254.1$, $p<.001$, $\eta2 = 0.92$. Accuracies were significantly increased the more trials were used to make the class predictions. However, there was no significant interaction between affective valence and number of trials, $F(8, 184) = 1.3$, n.s.

Cross-valence classification using workload conditions under positive affective valence as training data and workload conditions under negative affective valence as

test data resulted in an accuracy of 53.9 % (sd = 4.9 %), while classifier training on workload conditions under negative affective valence and classifier testing with workload conditions under positive affective valence yielded a mean accuracy of 54.0 % (sd = 5.6 %). Moreover, when using the workload conditions under neutral affective valence for classifier training, mean cross-valence classification accuracies were 54.3 % (sd = 5.3 %) for workload conditions under positive affective valence and 52.7 % (sd = 3.7 %) for workload conditions under negative affective valence. Finally, using the workload conditions under neutral affective valence for classifier training affective valence conditions under low WML (positive 1-back Vs negative 1-back) resulted in a mean accuracy of 52.2 % (sd = 5.0).

Discussion

Our investigation revealed that classification of WML under affective valence can lead to useful classification accuracies. Furthermore, we found that data integration over time yielded improved classification accuracies. However, resulting workload classifiers were context-specific and did not generalize to different affective contexts.

We found that classification of workload memory load under affective valence is possible, but only with diminished performance. Both, positive as well as negative affective valence resulted in decreased classification accuracies. Similar results have been found in a study by Mühl, Jeunet, & Lotte (2014), who used an N-back task to manipulate WML while social stress was induced with a stress-induction protocol, based on the Trier Social Stress Task (Kirschbaum, Pirke, & Hellhammer, 1993). The authors attempted to detect mental workload during stress using features from the frequency domain as well as the time domain. They concluded that it is possible to transfer classification methods across affective contexts, but only with diminished performance.

Mean classification accuracies found in our study exceeded the 70 % threshold commonly used for active BCIs (Kübler et al., 2001) already on the single-trial level and accuracies were further improved via integration over time to almost 90 % (87.7), which is comparable to previous research (e.g., Brouwer et al., 2012). Interestingly, our analyses revealed that data integration over two trials already resulted in the largest gains for classification accuracies. This indicates that workload classification

under affective valence can reach levels of accuracy that allow the use in natural environments.

As expected, we found that overall the frontal theta and parietal alpha features showed strong relevance for the distinction of WML. However, different EEG features were more or less relevant depending on the particular affective context. Activity in the alpha band appeared to be the most relevant feature for workload classification, especially under neutral affective valence (see Figure 26). Frontal theta activity appeared most relevant under negative affective valence, which might be due to the influence of frontal activity usually associated with affective valence reactions (see **Figure 27**). However, there were strong inter-individual differences in the data, which makes it difficult to draw strong conclusions. These strong inter-individual differences also highlight the necessity of individualized systems that are customized toward individual EEG patterns.

Cross-valence classification revealed that a classifier trained under low affective valence (neutral 1-back vs neutral 2-back conditions) was not sensitive toward changes in affective valence which did not involve changes in WML (positive 1-back vs negative 1-back). However, we also found that the resulting workload classifiers were highly context-specific, indicating that our classifiers cannot be used in different settings, which seems highly problematic of one wants to design brain-based human-machine systems that are robust to changes in the application setting. Surprisingly, these results are in conflict with results from the study by Mühl, Jeunet, & Lotte (2014), since their classifiers did generalize to different affective contexts. One reason for this conflicting results could be found in the use of additional frequency bands in their approach. We deliberately decided to exclude frequency bands not commonly associated with changes in WML to limit the effect of external factors (like muscular artifacts) on the results. Another reason for the different results might be the type of affect induction used in both studies. Our study used affective pictures as stimulus material, while the study by Mühl, Jeunet, & Lotte (2014) used a single change in external context (stress induction via performance evaluation) as affect induction.

Limitations and Outlook

EEG measures used to detect WML have shown to be sensitive to changes in affective valence. This failure to generate classifiers that generalize well under different affective

contexts raises various concerns. Future studies need to investigate if results generalize to different samples, since female subjects tend to show stronger reactions toward affective stimuli (Lang, Greenwald, Bradley, & Hamm, 1993) and also exhibit more stable responses (Ahern & Schwartz, 1985). Furthermore, it would be interesting to see if changes in stimulus material also impact our results. A study by Mühl et al. (2011) used visual, audio and audiovisual material for affect induction and found modality-specific responses in the EEG.

Finally, most EEG features used to infer the current user state suffer from many to many relations, meaning that different processes can influence the accuracy of mental state detections. Roy et al. (2013), for example, induced fatigue and found that with increasing time on task the discriminability of WML was decreased. Equivalent dipole fitting of independent components (Jung et al., 2001; Kothe & Makeig, 2011; Makeig, Bell., Jung, & Sejnowski, 1996; Makeig, Leslie, & Mullen, 2011; Oostenveld, Fries, Maris, & Schoffelen, 2011) has proven to be a useful approach that can be utilized to identify brain regions involved in the processing of different tasks. This method might therefore provide alternative measures that can be used to infer WML.

To cope with the failure to develop classifier models that are robust under different contexts, we either have to develop models that allow to account for such context-sensitivity, or find new, context-independent features. Otherwise, it seems, we will be forced to limit the use of such brain-adaptive systems to very specific application scenarios.

Acknowledgements

### 6.2.1 Additional Chapter: Valence Classification Results

Affective valence classification was based on the approach used for classification of WML, however only features in the alpha range (7-13 Hz) were used. Additionally, to

investigate if frontal asymmetries can be used to automatically detect changes in affective valence on the single subject level, I computed difference scores for all frontal channel pairs (AF4 – AF3, F6 – F5, F4 – F3, F2 – F1, FC6 – FC5, FC4 – FC3, FC2 – FC1) and used them as classification features in the FAA approach.

All mean classification accuracies in the normal approach were above seventy-five percent, even with single trial classification and almost reached ninety percent (89.4%) via data integration across five trials. Classification accuracies from the FAA approach almost passed the seventy percent threshold (68.4%) and data integration over time even resulted in accuracies as high as eighty percent (80.5%). Median values are shown via box plots in **Figure 28** (left).
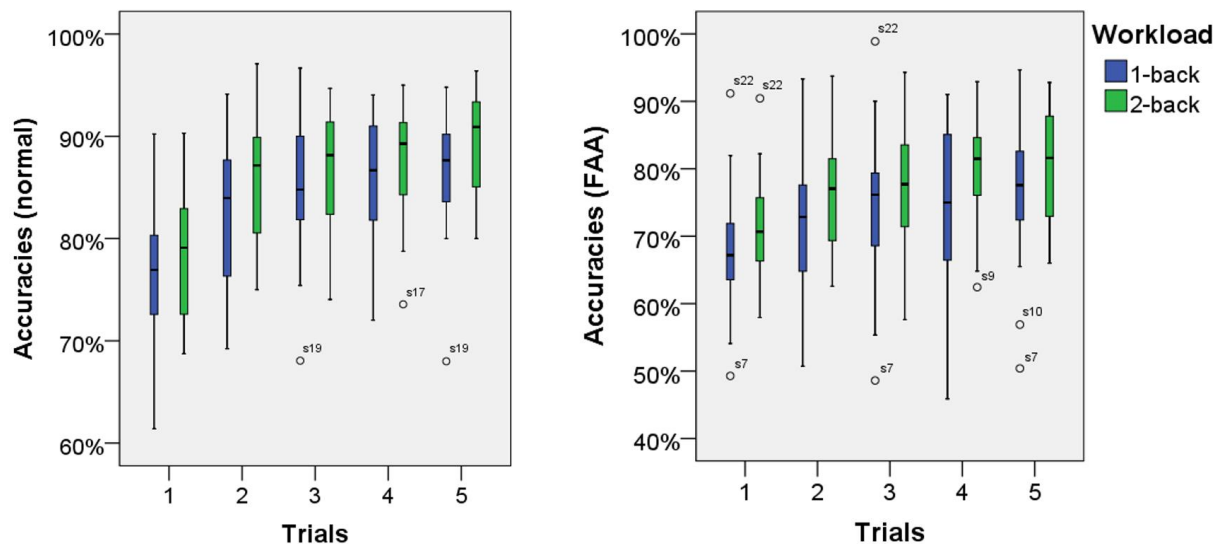
Figure 28. Box plots showing median values for accuracies for the normal approach (left) and the FAA approach (right): Blue boxes show accuracies for valence classification under low WML (1-back). Green boxes show accuracies for valence classification under high WML (2-back). Median values are indicated by black horizontal lines within the boxes. Top and bottom borders of the boxes represent the middle 50% of the data. Whiskers represent the smallest and largest values not classified as outliers (between 1.5 and 3 times the height of the boxes) or extreme values (more than three times the height of the boxes). Circles indicate outliers and stars show extreme values.

Surprisingly, for the discrimination of affective valence conditions under different WML, parietal alpha features were most important. However frontal alpha features were involved as well, as can be seen in the FAA approach (see **Figure 31** and **Figure 32**). Interestingly, discrimination of affective valence conditions under WML was mainly

based on features in the upper alpha range (see **Figure 29**, **Figure 30**, **Figure 31** and **Figure 32**).
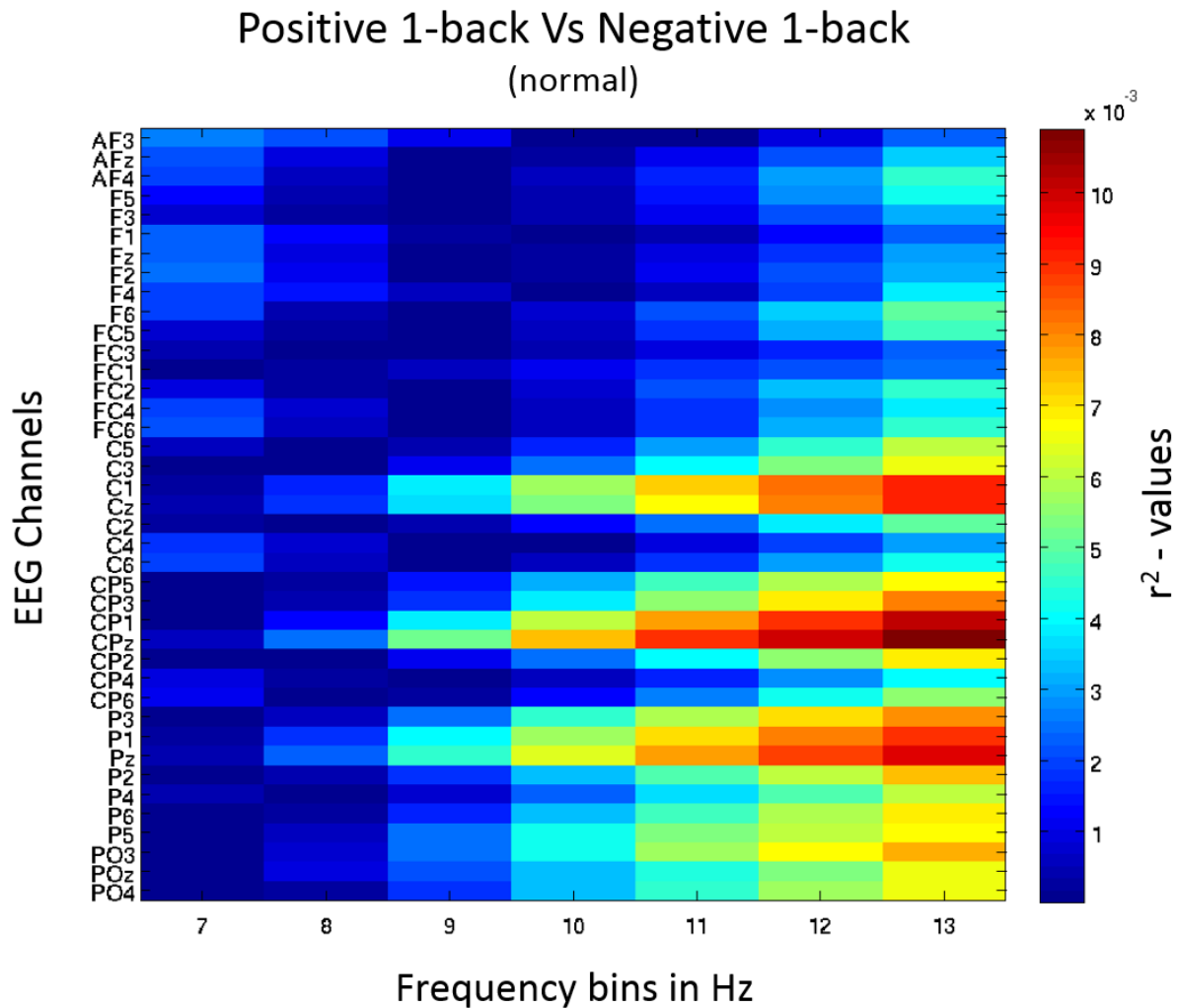
Figure 29. R² plot for the positive and negative affective valence conditions (normal approach) under low working memory load (1-back). Rows show individual EEG channels and Columns show individual frequency bins. R² values are color coded (see colorbar on the right side of figure for scaling).
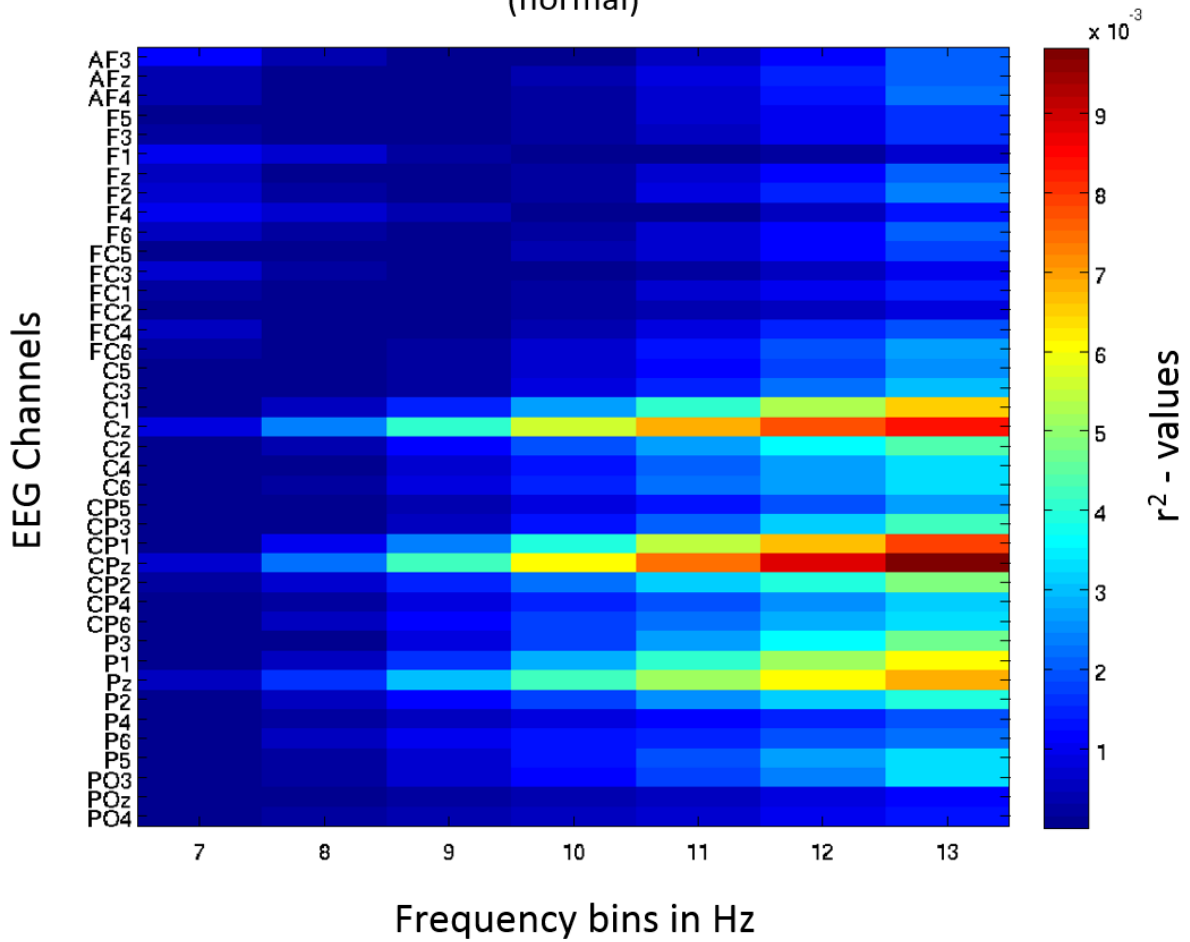
Figure 30. R² plot for the positive and negative affective valence conditions (normal approach) under high working memory load (2-back). Rows show individual EEG channels and Columns show individual frequency bins. R² values are color coded (see colorbar on the right side of figure for scaling).
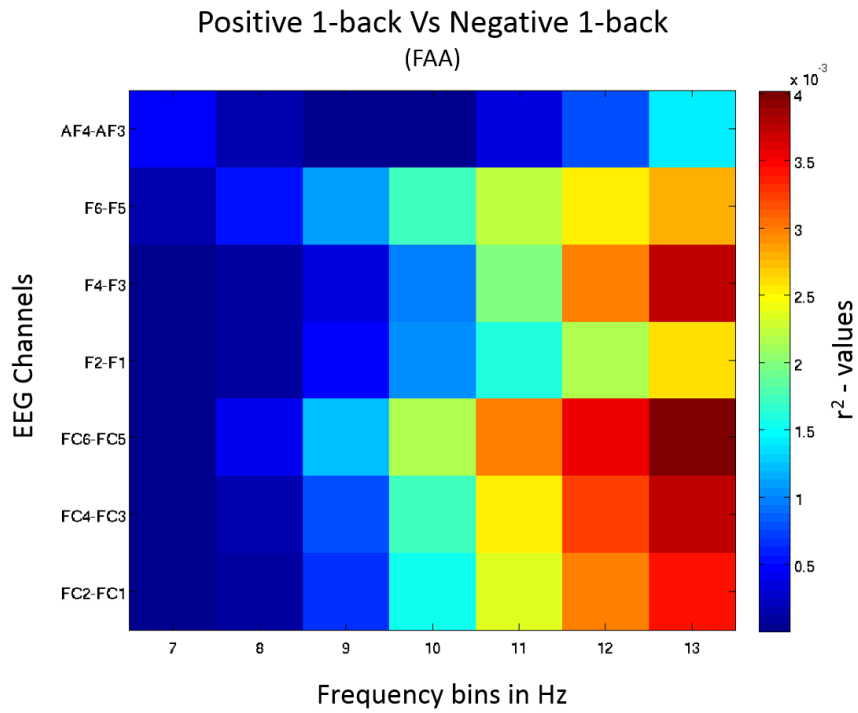
Figure 31. R² plot for the positive and negative affective valence conditions (FAA approach) under low working memory load (1-back). Rows show individual EEG channels and Columns show individual frequency bins. R² values are color coded (see colorbar on the right side of figure for scaling).
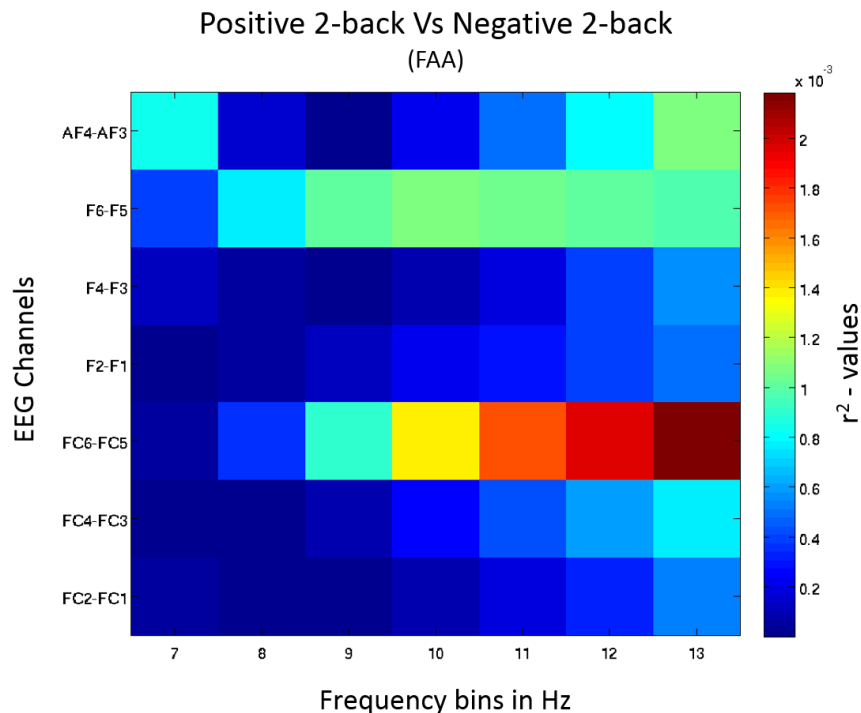
Figure 32. R² plot for the positive and negative affective valence conditions (FAA approach) under high working memory load (2-back). Rows show individual EEG channels and Columns show individual frequency bins. R² values are color coded (see colorbar on the right side of figure for scaling).

Interestingly, using classification accuracies in the normal approach as dependent variable there was a significant main effect for WML, $F(1, 23) = 7.1$, $p<.015$, $\eta^2 = 0.24$. Affective valence classification under high WML (2-back) resulted in higher classification accuracies than affective valence classification under low WML in the normal approach. As expected, we found a significant main effect for number of trials, $F(2.6, 59.2) = 96.4$, $p<.001$, $\eta2 = 0.81$. Accuracies were significantly increased the more trials were used to make the class predictions. Though, this effect was only present for up to three trials, after that data integration did not significantly improve classification accuracies. There was no significant interaction between WML and number of trials, $F(4, 92) = 0.9$, n.s. Median values are shown via box plots in **Figure 28** (left).

In the FAA approach there was also a significant main effect for WML showing increased classification accuracies under high workload, $F(1, 23) = 4.4$, $p<.05$, $\eta^2 = 0.16$. As expected, we also found a significant main effect for number of trials in the FAA approach, $F(2.8, 64.7) = 59.6$, $p<.001$, $\eta2 = 0.72$. Accuracies were significantly

increased the more trials were used to make the class predictions, but again, only up to three trials. There was also no significant interaction between WML and number of trials, $F_{(4, 92)} = 1.9$, n.s. Median values are shown via box plots in **Figure 28** (right).

Cross-workload classification using affective valence conditions under low WML (1-back) as training data and affective valence conditions under high WML (2-back) as test data resulted in an accuracy of sixty-four percent (sd = 7.8 %), while classifier training on affective valence conditions under high WML and classifier testing using affective valence conditions under low WML also yielded a mean accuracy of around sixty-four percent (64.8%; sd = 7.5%). Additionally, when using affective valence conditions under low WML for classifier training, mean cross-workload classification accuracies were around fifty percent (49.6%; sd = 4.5%) for workload conditions under neutral affective valence.

6.3 Brain-Computer Interfaces for Natural Environments require additional Context Information: An Illustration based on a Loss of Control Paradigm (LOC Study)

Sebastian Grissmann[1]*, Thorsten O. Zander[1,2], Josef Faller[3], Jonas Brönstrup[2], Augustin Kelava[1,4], Klaus Gramann[5,6] and Peter Gerjets[1,7]

[1] LEAD Graduate School, University of Tübingen, Tübingen, Germany

[2] Team PhyPA, Biological Psychology and Neuroergonomics, Berlin Institute of Technology, Berlin, Germany

[3] Laboratory for Intelligent Imaging and Neural Computing, Columbia University, New York, USA

[4] Hector Research Institute of Education Science and Psychology, University of Tübingen, Tübingen, Germany

[5] Biological Psychology and Neuroergonomics, Berlin Institute of Technology, Berlin, Germany

[6] Center for Advanced Neurological Engineering, University of California, San Diego

[7] Leibniz-Institut für Wissensmedien, Tübingen, Germany

Abstract

Most brain-computer interfaces (BCIs) focus on detecting single aspects of user states (e.g. motor imagery) in the electroencephalogram (EEG) in order to use these aspects as control input for external systems. This communication can be effective, but unaccounted mental processes can interfere with signals used for classification and thus impede BCI performance. To improve BCI performance, we propose deploying a heuristic approach that allows to describe different mental states influencing BCI performance. To test this approach, we analyzed neural signatures of potential affective states in data collected in a paradigm where the complex user state of perceived loss of control was induced. In this paper, source localization methods were used to identify brain dynamics with source located outside but affecting the signal of interest originating from the primary motor areas, pointing to interfering processes in the brain during natural human-machine interaction. In particular, we found affective correlates which were related to perceived loss of control. We conclude that additional context information may help to improve the applicability of BCIs to real-world scenarios.

Introduction

Traditionally, BCIs have been developed to translate measured brain activities into commands for technical systems in real-time. By using measures of neural activity like the electroencephalogram (EEG), BCIs therefore offer additional communication channels for human-machine interaction (Wolpaw et al., 2002). The primary goal in developing BCI systems was to support persons with severe behavioral impairments like amyotrophic lateral sclerosis (ALS, also known as Lou Gehrig's disease). However, recently there have been developments toward systems that can benefit users without disabilities as well.

From a user perspective, there are three main types of BCIs (Zander and Kothe, 2011): Firstly, reactive BCIs, tracking users` attention towards stimuli that are externally presented and encode direct control commands. Secondly, active BCIs relying on voluntarily induced changes in brain activity, e.g. by imagining motor activities. Both active and reactive BCIs are used to control devices through directed commands (e.g. P300 speller; Belitski, Farquhar, and Desain, 2011, or motor imagery based BCIs; Pfurtscheller and Neuper, 2006). Thirdly, passive BCIs apply mental state monitoring to establish an implicit control loop. Passive BCIs do not rely on a user`s awareness about the concrete information exchanged during this interaction. An example can be found in Zander et al. (2014), where ERP-based error detection was used to improve cursor movements toward targets in a 2d-grid. The participants of the study did not know, that they provided information to the system, but still recognized that system performance was improved. Other possible application scenarios for passive BCIs are adaptive learning environments (Gerjets et al., 2014), automatic correction of errors (Parra et al., 2003) or even the detection of mental states like bluffing in a game context (Reissland and Zander, 2009).

The Relevance of Complex Mental User States to BCI Performance: The Case of Loss of Control

Most available BCIs try to identify specific, narrow mental states or processes, as for example motor imagery (Lotte et al., 2010), working-memory load (Gerjets et al., 2014) or affective responses (Mühl, Heylen, et al., 2014). However, such systems typically do not account for contextual fluctuations in other mental states and processes that

arise during use. In line with this reasoning, previous work suggests that spontaneously occurring decrements in BCI reliability could be related to modulations in unaccounted mental processes and states (Zander and Jatzev, 2012; Myrden and Chau, 2015).

Natural environments for innovative BCI applications (e.g. classroom or workplace settings; cf. Gerjets et al. 2014) usually contain various factors that can potentially evoke complex mental states in a user. Therefore, such application scenarios might easily induce a variability of different mental states that are not directly related to the core mental state or process that is in the focus of a certain BCI application. Accordingly, such factors can produce changes in EEG signals that are picked up by a system, but cannot be interpreted correctly due to the limited framework (usually consisting of a single mental state or process) the system is built upon.

As many available classification algorithms assume stationary signal properties, which are learned from a single calibration session, these algorithms are unable to track shifts in the feature space (called non-stationarities) and rather require users to adapt to unexpected or erroneous behavior of the system, for instance by trying to produce brain signals similar to those used during calibration (Krauledat, Tangermann, Blankertz, and Müller, 2006).

Future systems should automatically adapt to contextual changes in the EEG and provide stable performance even in noisy environments outside the lab. To facilitate the development of such systems it is necessary to investigate complex user states with the aim to separate contextual and potentially interfering states from the primary interaction mode.

In this paper, we address the role of interfering mental states on BCI performance in a paradigm that was designed to experimentally manipulate an important complex user state in the context of BCI systems, namely the perceived feeling of a loss of control (LOC). The LOC paradigm used in our study has already been hypothesized to evoke affective as well as cognitive responses (Zander & Jatzev, 2009) and thereby appears to be particularly suited to study different basic components of complex user states. In this paradigm, the feeling of LOC is evoked by means of a feedback manipulation: Subjects were first trained to use a fixed set of rules (color-angle associations) to control a simple letter rotation task by means of button presses with the left versus right hand. Later in the experiment, the previously learned rules

were temporarily violated, thus eliciting the feeling of LOC. This paradigm allows to study the neural signatures of motor executions with and without LOC in order to analyze what happens to a specific BCI signal of interest (in this case motor execution responses) when other cognitive or affective mental states overlap with this signal.

Contrary to standard BCI approaches, we decided to analyze motor execution responses rather than motor imagery responses in order to better control for the perceived LOC. Input based on motor imagery can be rather unreliable, whereas button presses are an extremely accurate input modality thereby not yielding unintended LOC experiences. Moreover, earlier work by Pfurtscheller and Neuper (1997) has demonstrated that that motor execution responses are structurally similar to motor imagery responses with respect to their neural signature. Therefore, for the purpose of the current study, namely to investigate the role of interfering mental states on the stability of BCI performance, we assumed that using a button-press paradigm would yield similar results than using a paradigm based solely on motor imagery.

In earlier work it has already been demonstrated that EEG signature of motor execution responses changed under LOC, thereby leading to a BCI performance degradation when the interface was controlled by EEG signals (Zander and Jatzev, 2012). However, in this work it remained unclear, which aspects of the complex user state contributed to the altered signature of motor responses in the EEG signal.

Following previous work by Zander and Jatzev (2009), we assume in this paper that LOC might involve affective processes. According to this reasoning, the feeling of losing control over an interface should provoke negative emotions like irritation, worry, frustration, anger or helplessness (Reuderink et al. 2009). The aim of the current paper is to provide additional evidence for this assumption by comparing LOC trials to correct trials in our experimental paradigm with regard to potential neural signatures of affective user states. For this purpose we will not only analyze sensor-based EEG measures but also conduct an EEG-based source analysis.

Neural Signatures of potential Affective User States in the Loss of Control Paradigm

Several studies used EEG measures to investigate the neural basis of affective processes (Davidson, 1992; Kim et al., 2013; Olofsson et al., 2008). Laterality in the alpha band measured in frontal regions, also called Frontal alpha asymmetry (FAA), is

widely used in these studies as a measure for affective processing. However, this measure seems to capture a wide variety of affective responses. Many studies that used brain lateralization as a measure of affective states or responses focused on one particular dimension of emotional experience, namely the valence dimension (Ahern & Schwartz, 1985; Davidson & Tomarken, 1989; Huang et al., 2012; Lin et al., 2008; Ramirez & Vamvakousis, 2012). The valence dimension reflects the pleasantness or unpleasantness of an affective event and ranges from extreme pain (or unhappiness) to ecstasy (or extreme happiness). Left frontal activity in this context often indicates a tendency toward positive emotions like happiness (e.g. Ahern & Schwartz, 1985).

However, the equivalence of left frontal activity with positive emotions is not unequivocal. Several studies showed that some emotions with negative valence might also activate the left hemisphere. An interpretation of this apparent contradiction has been provided by the Approach-Withdrawal model, first proposed by Schneirla (Schneirla, 1959, as cited in Dalgleish, 2004). It states that the left hemisphere is active in emotions associated with pleasant or unpleasant approach behaviors like anger or engagement whereas the right hemisphere is related to avoidance behaviors like fear. As an alternative explanation it has been proposed that FAA can be used to measure emotional responses along the dominance dimension (Demaree, Everhart, Youngstrom, and Harrison, 2005). Dominance has been specified as 'feelings of control and influence over everyday situations, events and relationships versus feelings of being controlled and influenced by circumstances and others' (Mehrabian, 1994). In this context left-frontal activation is related to more dominant emotions that can also be of positive and negative valence. Whichever of these interpretations is correct, FAA should in any case be sensitive to affective responses we expect to be modulated during perceived loss of control. Interestingly, a recent study by Reuderink, Mühl, and Poel (2013) found that activity in the lower alpha band was related to the valence dimension, while activity in the upper alpha band was related to the dominance dimension, which might be of important in this analysis, since it seems plausible that LOC might have an effect on the dominance as well as the valence dimension.

Thus, in a first step to explore whether the complex LOC user state contains affective components besides motor activities, we focused on examining a potential FAA shifts occurring during phases of reduced control. Specifically, using the FAA

measure, we investigated each individual experience of LOC during phases with loss of control (as compared to trials within these phases that do not reflect a LOC).

In a second step, to further substantiate potential affective correlates found using the FAA, we conducted an additional cluster analysis to explore whether the corresponding brain dynamics can be attributed to specific neuroanatomical structures. In order to provide evidence for specific structures as sources of a FAA we used equivalent dipole modeling of independent components derived from an independent component analysis (ICA; Makeig et al., 1996; Oostenveld & Oostendorp, 2002). However, due to the novelty of the experimental paradigm used it is difficult to make concise predictions about specific brain areas involved, especially since the previous literature on brain laterality concluded that most theories underlying frontal asymmetries lack anatomical specifity (Wager et al., 2003). Nonetheless, potentially revealed brain sources of a FAA are expected to be interpretable as affective structures in the context of the paradigm used in this study.

Material and Methods

Experimental Task

A simple game (called 'Rotation-Left-Right or RLR') was used to modulate LOC over the course of the experiment. Subjects had to discretely rotate a letter stimulus until it was aligned to a target letter with regard to its rotation angle. The basic principles of the task are illustrated in **Figure 33** and a full description of the experimental task can be found in Zander and Jatzev (2012).

Stimulus display

Rotating figure

Target position

Letter **R**: Clockwise rotation
with right hand
Letter **L**: Counterclockwise
rotation with left hand

Experimental conditions

Normal

Letter color
indicates
rotation angle

e.g. green = 30°

e.g. red = 90°

Loss of control

**Color mapping
is changed!**

LoC occurs
randomly with a
maximal rate of
30%

e.g. red ≠ 90° → **30°**
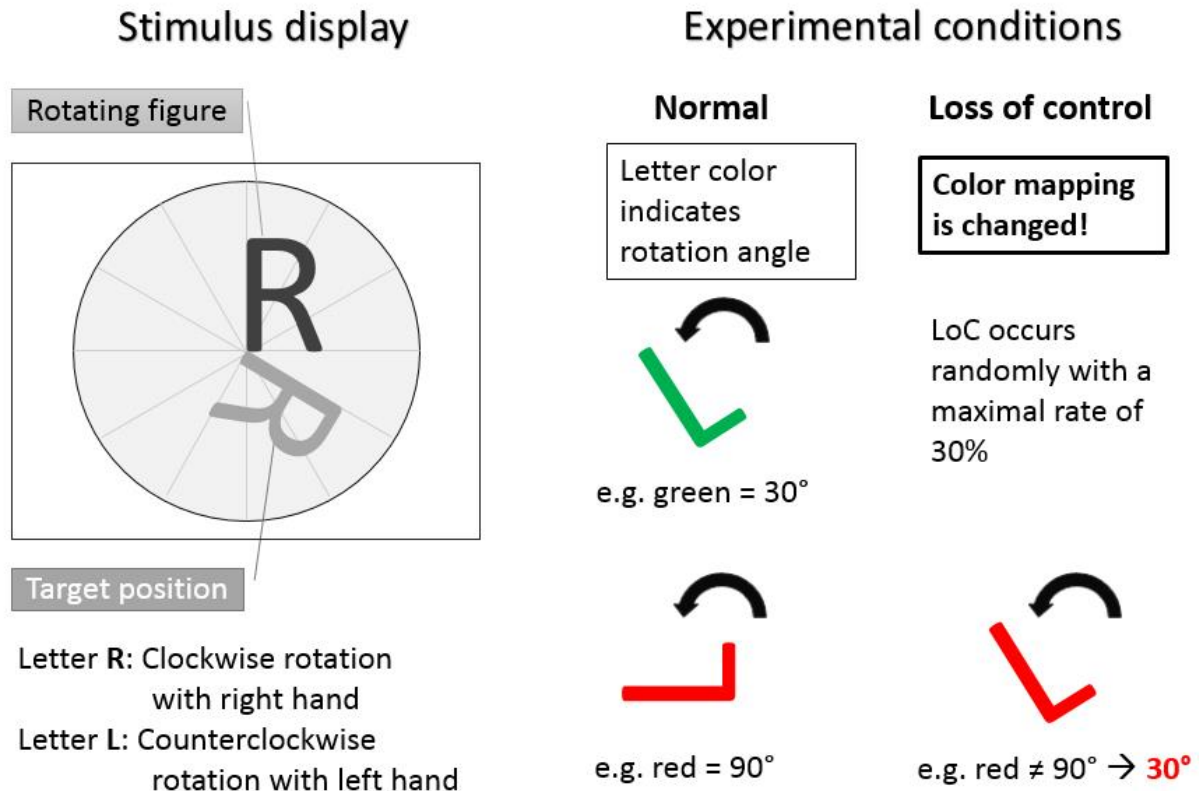
Figure 33. The Rotation-Left-Right (RLR) game used as experimental task. Left: Stimulus display showing stimulus to be rotated and target position. Letter R indicated clockwise rotation with right hand button press. Letter L indicated anti-clockwise rotation with left hand button press. **Right:** Experimental conditions. In the normal trials condition the color-rotation angle mapping was fixed. In the loss of control (LOC) trials condition was distorted with a maximal occurrence rate of 30%. Hence, during LOC the subjects did not know how the system responded to a button press. Adapted from Zander (2011)

The stimuli were the two capital letters R and L. The capital letter R indicated clockwise rotations after a button press of the right hand and the capital letter L indicated counterclockwise rotations after a button press of the left hand.

The color of the stimulus changed to indicate an expected rotation angle with the color being either, green for 30, yellow for 60 or red for 90 degrees. For each stimulus rotation, subjects could decide whether they wanted to rotate the letter stimulus according to the angle indicated by the color by pressing a button or not. A training phase was provided to allow subjects to develop better strategies to finish the task successfully in a shorter period of time. Usually, better strategies resulted from waiting for a rotation opportunity with a larger rotation angle. All colors and corresponding rotation angles occurred with the same probability and each stimulus had to be rotated for at least 90 degrees in order to be aligned with a target letter. In

the first part of the experiment, the contingencies between letters, colors and rotations were trained for approximately 30 minutes. In the second part of the experiment the previously learned rule system of color-angle mapping was systematically violated by rotating the stimulus by a different than expected angle with a chance of up to 30%. These violations resulted in trials with unexpected rotations (LOC trials), forced subjects to reconsider and potentially modify their previously formed strategies and induced – according to our hypothesis – a complex user state, namely the feeling of a perceived LOC.

Experimental Setup

The Dataset

The dataset comprised eighteen healthy subjects (age range: 19 to 40 years) that were recorded in a previous experiment to investigate the effect of perceived LOC on EEG features related to motor response (Jatzev et al., 2008). The recordings have been conducted in line with the Declaration of Helsinki. Out of these eighteen subjects, three subjects had to be removed from the analysis due to strong artifact contamination. EEG-data was recorded with 32 active Ag/Cl-electrodes (positioned according to the extended 10% system) and a biosignal amplifier (Brainproducts GmbH., Gilching, Germany). All channels were referenced to the nasion. Impedances were kept below 20 kΩ. The data were sampled at 1 kHz. The EEG recording sessions lasted for about sixty-six minutes. Our analysis was focused on data from the second part of the experiment where trials with expected stimulus rotation (normal trials) and trials with erroneous stimulus rotation (LOC trials) were mixed. At the beginning of the analyzed data segment with an overall duration of fifteen minutes, the incidence of trials where previously learned rules were violated was linearly increased from 0 to 30% over four minutes. Then the occurrence of LOC trials stayed at a maximum of 30% for 7 minutes, before the incidence of trials with incorrect rotation (LOC trials) was linearly decreased over the next 4 minutes. The experiment ended with 7 minutes consisting only of normal trials.

Preprocessing

First, the data was bandpass filtered between 1 and 30 Hz. Then it was down-sampled to 250 Hz and re-referenced to the common average (CAR) to suppress artifactual activity that spread over all channels (e.g. muscle activity or outside electrical noise;

see McFarland, McCane, David, and Wolpaw, 1997). To ensure that our findings could be applied to an online application scenario no further manual artifact rejection was performed. A logarithmic transformation was applied before computing the spectra to better fit the data to a Gaussian distribution. Based on the literature and the available data, the electrode pairs F3 / F4 and FC1 / FC2, respectively, were analyzed with regard to potential FAAs. All data processing was done via MATLAB and Statistics Toolbox Release 2013b (The MathWorks, Inc., Natick, Massachusetts, United States) and EEGLAB (Delorme & Makeig, 2004).

Analysis

Each rotation trial started with the presentation of a letter stimulus (L or R, either colored in green, yellow or red) together with a target letter in grey to which the stimulus letter had to be aligned to by means of an appropriate rotation. Subsequently, subjects had a time window of 1000 ms to react (or not) with a left hand (for the letter L) or right hand (for the letter R) button press, resulting in a stimulus rotation (or not). After 300 ms of the subsequent 850 – 950 ms interstimulus interval the color of the letter stimulus turned to grey. Following the interstimulus interval, a new letter stimulus was presented in color to start the next rotation trial. Only rotation trials with button press were used for analysis. The time window for analysis started 100 ms after the rotation of the stimulus and ended 700 ms later (50 – 150 ms before the interstimulus interval ended) to ensure that the data was not contaminated by effects related to expectancy of the next rotation opportunity (see **Figure 34**, left).
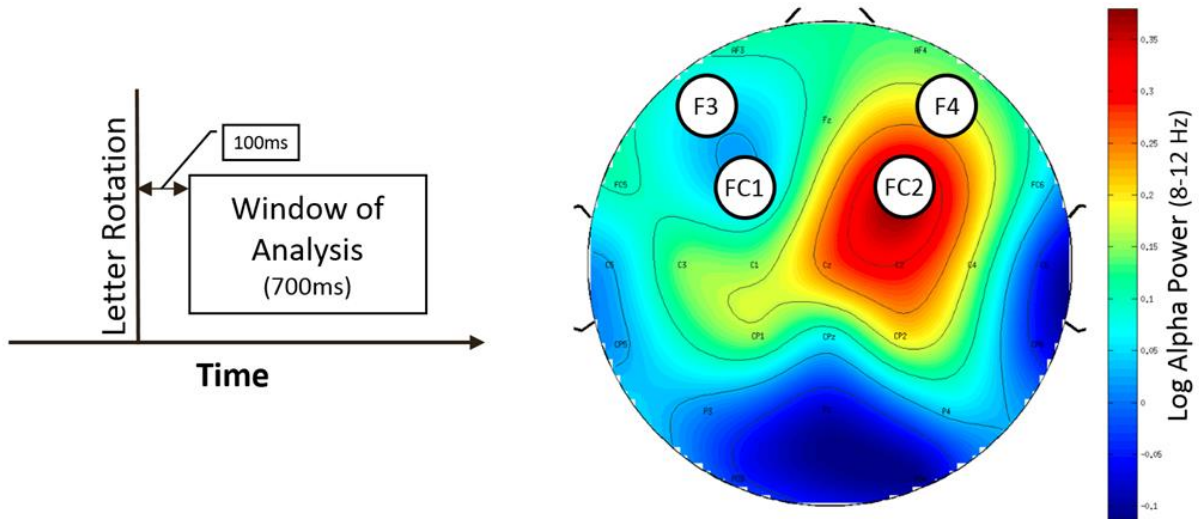
Figure 34. Window of analysis (left) and topografic difference plot illustrating the difference in alpha power between normal and LOC trials (right). Left: Time is on the horizontal axis. Vertical line indicates onset of letter rotation. **Right:** Topographic difference plot showing frontal alpha (8-12 Hz) asymmetry effect. Nose is at the top. Channel labels indicate channel location. Electrodes of interest are marked with black circles. Colorbar indicates value range of alpha power (after logarithm). Blue color shows low alpha power (top left and bottom left). Red color shows high alpha power.

Power spectral density (PSD) of the EEG signal was used as basis for FAA estimation between 8 and 10 Hz for the lower alpha band and between 10 and 12 Hz for the upper alpha band. FAAs were computed as difference score (ln(Right alpha power) − ln(Left alpha power)), following the approach from Allen, Coan, and Nazarian (2004).

For the analysis, we performed a three-way repeated measures ANOVA using the grand average (summing across all trials) FAA as dependent variable. The first factor was trial type with two levels (normal and LOC). The second factor was electrode pair, also with two levels (F4 – F3 and FC2 – FC1). The third and last factor was alpha band (lower alpha = 8 to 10 Hz and upper alpha = 10 – 12). Effects only appearing in the data of individual subjects were not taken into account.

For the cluster analysis, equivalent dipole modeling based on topographies of independent components (ICs) was used to estimate the location of cortical sources underlying potential effects at the sensor level. Dipole fitting was done using a simple three-shell spherical head model implemented in the DIPFIT-plugin from the EEGLAB toolbox (Delorme et al., 2011; Oostenveld, Fries, Maris, and Schoffelen, 2011). To

increase location accuracy of the resulting component clusters, all independent components that contained dipole coordinates with residual variance greater than 15 % have been automatically removed before clustering. Component clustering was done based on mean IC log spectra, equivalent dipole locations as well as scalp maps as determining factors. Principal component analysis (PCA; Law & Jolliffe, 1987) was used to compress these measures (excluding dipole locations) to their first 10 principal components. The dipole locations were naturally three-dimensional. To extract closely spaced component cluster these three-dimensional dipole locations were weighted by a factor of 20. All these measures were finally compacted into 10 dimensions via PCA, since clustering algorithms may not work well with measures having more than 10 dimensions. K-means clustering (Ding & He, 2004) was used and any outlier ICs were automatically removed if the distance to any cluster centroid in joint measure space was greater than three standard deviations from the mean. The resulting IC clusters were inspected and ICs that exhibited activity that seemed to originate from non-brain artifacts were removed based on their activation spectra, scalp topographies and dipole location. Anatomical regions were assigned to the resulting component cluster centroids using the Talairach web client (Lancaster et al., 2000). Only clusters exhibiting a significant (p<.05) difference in the alpha band (8-12 Hz) between conditions were used for analysis.

Results

Scalp-based Analysis:

As expected, we found a main effect for trial type, $F(1, 14) = 9.46$, p<.01, $\eta^2 = 0.40$. LOC trials showed stronger left frontal activation (due to higher right frontal alpha power) than normal trials. There was no significant main effect for electrode pair, $F(1, 14) = 2.15$, n.s. Furthermore, there was no significant main effect for alpha band, $F(1, 14) = 0.65$, n.s. Finally, none of the interactions exhibited significant effects (all F-values < 1.5). A topographic difference plot that illustrates the effect across both alpha bands at the electrode level can be seen in **Figure 34** (right side), median values are shown in **Figure 35** and channel spectra for all electrodes are shown in **Figure 36**.
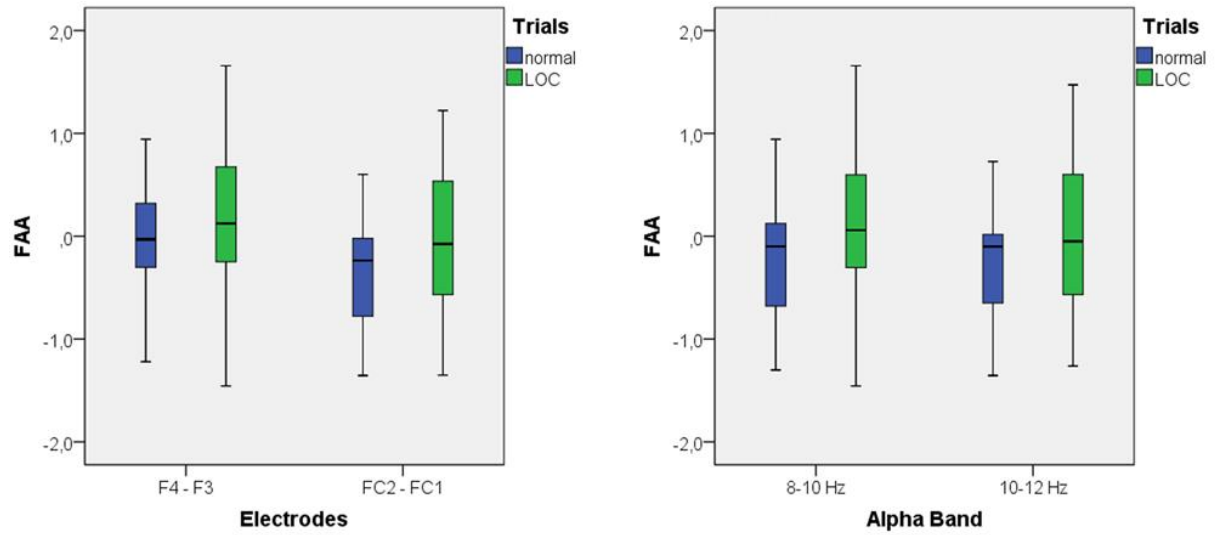
Figure 35. Box plots showing median values for electrodes (left) and alpha bands (right).
Median values are indicated by black horizontal lines within the boxes. Top and bottom borders of the boxes represent the middle 50% of the data. Whiskers represent the smallest and largest values not classified as outliers or extreme values.
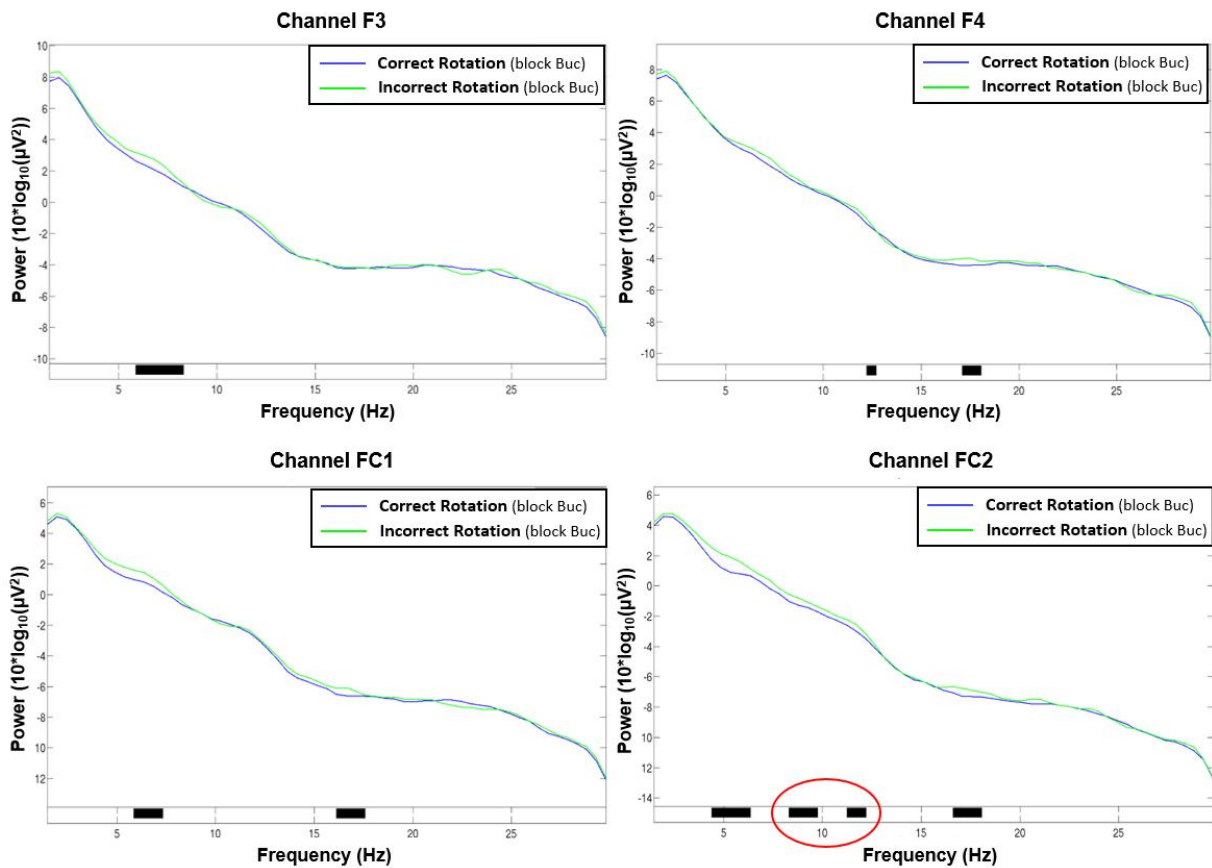
Figure 36. Channel spectra between 1 to 30 Hz for all four channels used in the analysis.
Top left: Channel F3. Top right: Channel F4. Bottom left: Channel FC1. Bottom right: Channel FC2. Significant differences (p<.05) are indicated with black bars. Please note that only channel FC2 shows significant activity in the alpha range (8-12 Hz; marked in red).

Cluster Analysis: Identifying Sources underlying the FAAs found at the Sensor Level

The cluster analysis yielded an IC cluster that showed spectral activity likely related to frontal laterality responses. The spectral activity of the component cluster projected strongly onto an area close to electrode FC2 (see **Figure 37**, bottom left). The scalp projection of this cluster showed the highest correlation (r=0.64, p <.001, sig.; see **Table 1**) with the topographical difference plot (see **Figure 34**, right), when compared with all other component clusters that exhibited a significant effect in the alpha band. Hence, the topographies show a unique, high similarity.

Table 1. Correlation coefficients between difference plot and all component clusters with significant activity in the alpha band. Please note that the right BA6 cluster displays the highest correlation coefficient (marked in red), indicating the largest overlap between the component cluster and the difference plot.

Cluster	right BA39	left BA6	right BA6	right BA10	right BA37
ρ	-0,24	0,17	**0,64**	0,38	-0,43
All correlations significant at p<.001					

This component cluster was located in or near the right supplementary motor area (SMA) in Brodmann Area 6 (BA6; see **Figure 37**) and showed decreased lower alpha activity (p <.05) and increased upper alpha activity (p <.05) during LOC trials.
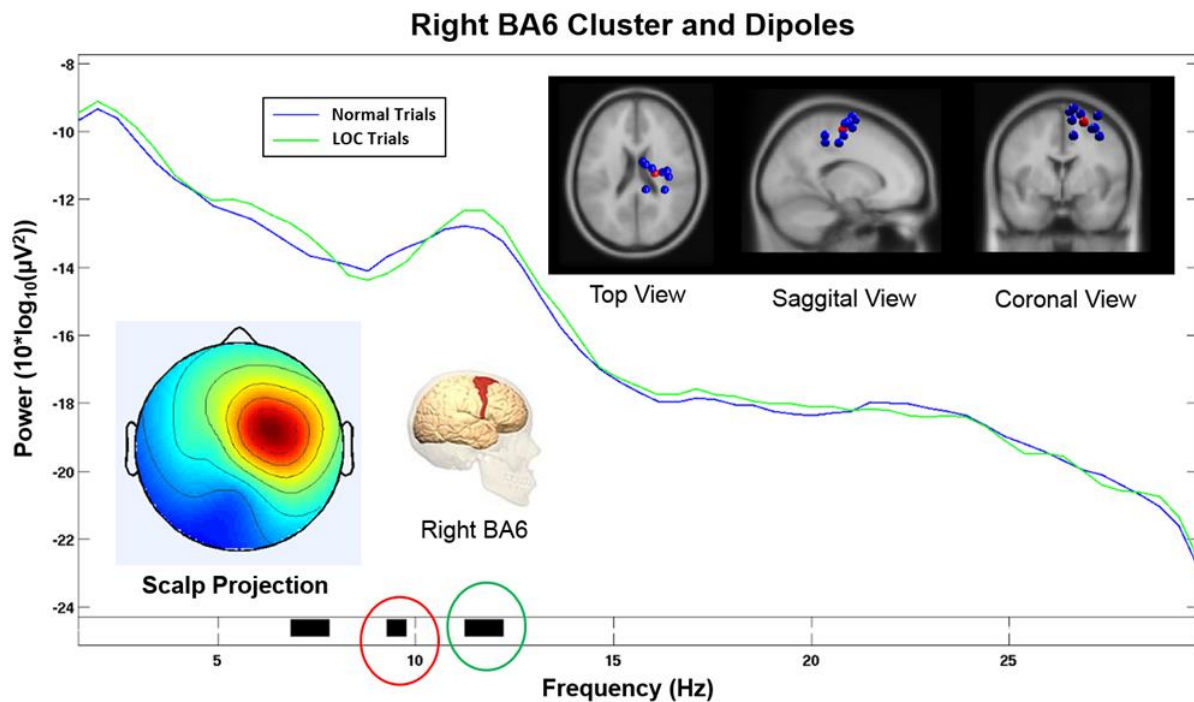
Figure 37. Component cluster in or near right Brodmann Area 6 (BA6). The spectrum shows a significant decrease (p<.05) in the lower alpha range (black horizontal bar marked in red) and a significant increase (p<.05) in the upper alpha range (black horizontal bar marked in green) during LOC. Top right shows the location of individual dipoles as well as the cluster centroid (red). Bottom left shows the corresponding scalp projection (red colors indicate higher cluster weights). Bottom middle shows the right BA6 as a whole (image taken from https://www.wikimedia.org).
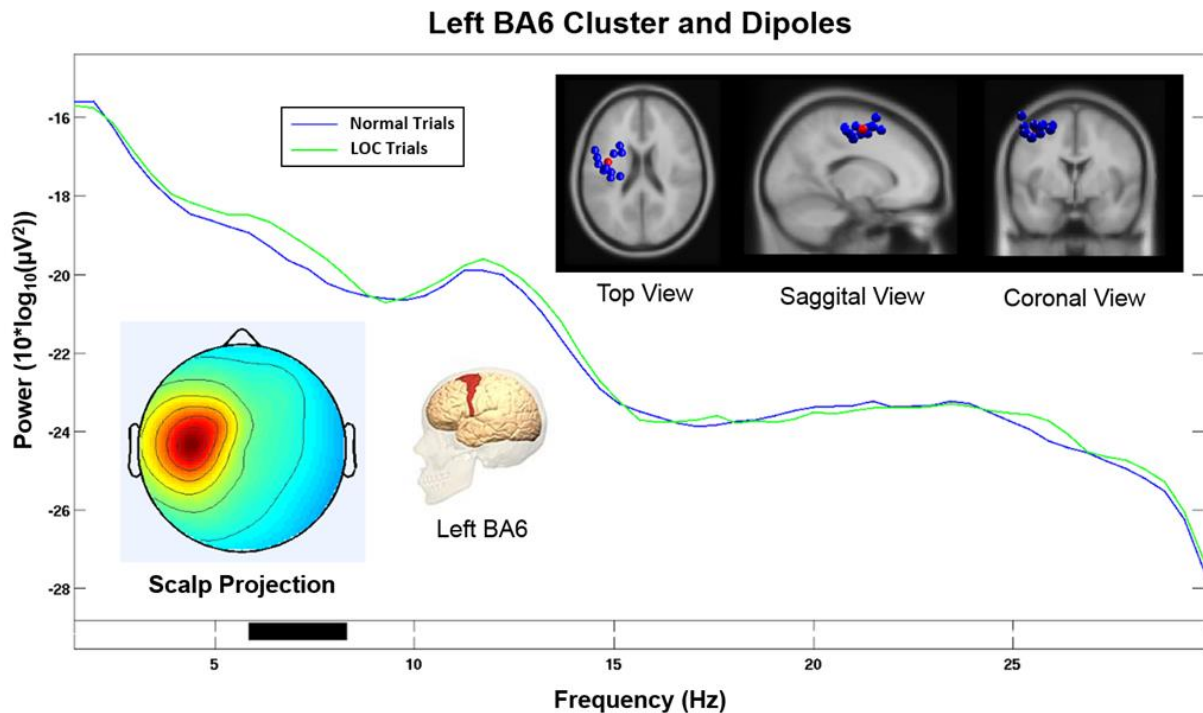
Figure 38. Component cluster in or near left Brodmann Area 6 (BA6). Note that the spectrum does not show a significant increase (p<0.05) in the alpha range during LOC. Top right shows the location of individual dipoles as well as the cluster centroid (red). Bottom left shows the corresponding scalp projection. Bottom middle shows the left BA6 as a whole (image taken from https://www.wikimedia.org).

The homologous cluster located in or near the left BA6 (see **Figure 38**) did not show significantly increased alpha power during LOC. A visual inspection of the activation spectra of both clusters did not show any indication of electromyographic contamination.

Discussion

We could demonstrate that the complex user state in the focus of this paper, namely the feeling of perceived loss of control, displays some of the characteristic neural signatures of affective processes. In particular, our scalp based analysis revealed a frontal asymmetry in the alpha band, indicating an increase in left frontal activity (due to an increase in right frontal alpha power) for incorrect stimulus rotation (LOC trials) as compared to normal trials with correct stimulus rotation. We further approximated the cortical sources that seem to contribute to the LOC-related FAA using equivalent

dipole fitting of independent components, revealing opposing effects that could not be found at the electrode level.

Frontal Alpha Asymmetry indicates affective Responses during perceived Loss of Control

We found that trials with incorrect stimulus rotation (LOC) were accompanied with increased left frontal activity (due to increased right frontal alpha power). Such responses are commonly related to increased dominance, emotions associated with approach behaviors, as well as positive valence. While a shift toward positive valence would seem quite puzzling, an emotional shift toward more dominance and approach behaviors could be well explained with feelings of anger, hostility and contempt (Demaree et al., 2005), which possibly were experienced by subjects during LOC trials. These approach related responses would reflect engagement (Harmon-Jones, Gable, and Peterson, 2010), indicating that subjects made efforts to regain control of the system. This interpretation appears most likely, especially since no behavioral alternative was available (subjects could only press one button), previously formed strategies to solve the task did not work anymore due to the violation of task rules and the random nature of the occurring LOC trials made it difficult to adapt those strategies. Interestingly, the fact that the system did not respond properly anymore, impeding task performance, did not seem to result in task disengagement, since task disengagement should have been reflected in decreased left frontal activation, thereby implying avoidance behaviors resulting in feelings like frustration.

Component Cluster Analysis

Scalp electrodes record a mixed sum of cortical activity which can distort results and lead to incorrect conclusions. Independent component analysis uses a linear transformation of the scalp signal to separate the recorded activity into different data sources. In combination with equivalent dipole fitting it is possible to investigate the cortical dynamics that underlie the signals recorded at the sensor level. We found a component cluster that was located in or near BA6 displaying a scalp projection (see **Figure 37**, bottom left) closely resembling the laterality effect found in the alpha power difference plots (see **Figure 34**, right). This cluster exhibited an alpha band activity during LOC trials that was decreased in the lower alpha band (p<.05) and increased in the upper alpha band (p<.05; see **Figure 37**). Similar results have been found in a

study by Reuderink, Mühl, and Poel (2013). The authors found that activity in the lower alpha band was related to the valence dimension, while activity in the upper alpha band was related to the dominance dimension. However, our investigation did not exhibit this effect at the electrode level which is the basis of most other FAA related analyses. The reason for this discrepancy is probably due to the fact that scalp electrodes record a mixture of multiple different cortical sources, which can distort effects based on single cortical sources. However, the sample was too small to put a strong emphasis on the results of the cluster analysis and further studies are needed to support these findings. Nevertheless, we see this as further evidence that that the FAA effect might be more complex than previously assumed. A separate analysis of different alpha bands in combination with source based analyses might allow to disentangle different responses that are usually associated with this measure.

The component cluster that we found to be related to the FAA induced by LOC was located in or near BA6, which plays a plays a major role in the planning of motor responses. In line with this function of BA6, previous research has indicated that emotional experiences are always associated with certain motor response tendencies (e.g., Önal-Hartmann, Pauli, Ocklenburg, and Güntürkün, 2012), for instance, we reach for positive stimuli or push away negative ones. BA6 is composed of the premotor cortex and, medially, the supplementary motor area (SMA). The right BA6 cluster centroid was located closest to the SMA (see **Figure 37**, top right). The SMA has been shown to be active during motor action under affective influence. For example, SMA activation was found for emotional conflict in a study by Ochsner, Hughes, Robertson, Cooper, and Gabrieli (2009) in an affective flanker task. Another study by Oliveri et al. (2003) found enhanced motor evoked potentials in the SMA when using emotional stimuli. The authors conclude that the SMA plays a role when emotional experiences are transformed into motor actions. This evidence is in line with our expectations concerning the emotional effects of the paradigm used in our study.

Our findings are further supported by the fact that the homologous component cluster located on the left hemisphere and projected onto electrodes F3 and FC1 (see **Figure 38**) did not show a significant increase in the alpha band during LOC trials. Furthermore, the scalp projection of this component cluster did not exhibit a strong similarity (r=0.17, p <.001, sig.; see Table 1) with the topographical difference plot (see

Figure 34, right). We therefore conclude that the right BA6 component cluster is the main contributor to the laterality effect (measured via FAA) we found at the sensor level.

Limitations and Outlook

Although we were able to identify affective responses in the LOC task as expected, this work has two main limitations. First, no classification has been conducted on the discovered laterality response. Future studies would need to adapt the LOC paradigm to make it suitable for a real-time classification approach. Such a paradigm should use motor imagery instead of button presses as input and implement a passive BCI for the automatic detection of affective processes that might interfere with the primary interaction mode. However, building such a system is not trivial, since both the classification of motor imagery as control input as well as the detection of affective responses will not be 100% reliable and a separation of both influences on the scalp recorded EEG might prove difficult. Using ICs as classifier features might turn out to be a viable approach, but up to now it is quite unclear whether IC cluster remain stable when using them in such a dynamic system.

Second, future studies should try to obtain more online measures of the subjective user states during interaction with the system. Since classification is done on a trial by trial basis, non-intrusive approaches would be ideal (e.g., heart rate or electrodermal activity). However, such approaches pose their own difficulties and usually do not work perfectly (Mauss & Robinson, 2009). Due to the necessary temporal density of user state measures, subjective ratings that can only be acquired in retrospect do also not seem very suitable for this paradigm. Nevertheless, they might provide an approximation of the general experience during the experiment and it is strongly recommended that future studies should implement them (e.g. affective ratings for all three dimensions from the PAD model using the self-assessment manikin; Bradley and Lang, 1994). With such measures it might become possible to uncover inter-individual differences, thereby enabling a more differentiated perspective. For instance, we assume, that even if most subjects were not discouraged by the erratic behavior the system exhibited, some subjects might have been.

Furthermore, the current paradigm does not manipulate cognitive and affective processes separately and therefore these processes are confounded to a certain

degree. Future studies should use paradigms that allow for such a manipulation to further disentangle the specific brain responses related to cognitive as well as affective responses. Potential candidates are the affective flanker task (Alguacil, Tudela, and Ruz, 2013) or the affective n-back task (Passarotti, Sweeney, and Pavuluri, 2011).

Conclusion

Using FAA in combination with source localization approaches, we described brain responses in the EEG during naturalistic human-machine interaction that appear to be related to affective components of a complex user state (LOC). Our analyses helped to distinguish these responses from those related to the primary interaction mode (button press).

This is further evidence that complex user states involve multiple factors. Such factors might be used as additional context information in future BCI systems by extracting a multi-facetted framework from the ongoing EEG. Such a framework might help to increase reliability in uncontrolled settings by providing information which is usually hard to access.

Acknowledgments

6.3.1 Additional Section analyzing Activity of the right BA6 Cluster

Additional analysis of activity of the right BA6 cluster over the course of the second half of the experiment revealed that this cluster exhibited increased alpha activity at the end of the experiment (normal condition after LOC) when both trial types (normal and LOC) were combined (see **Figure 39**).
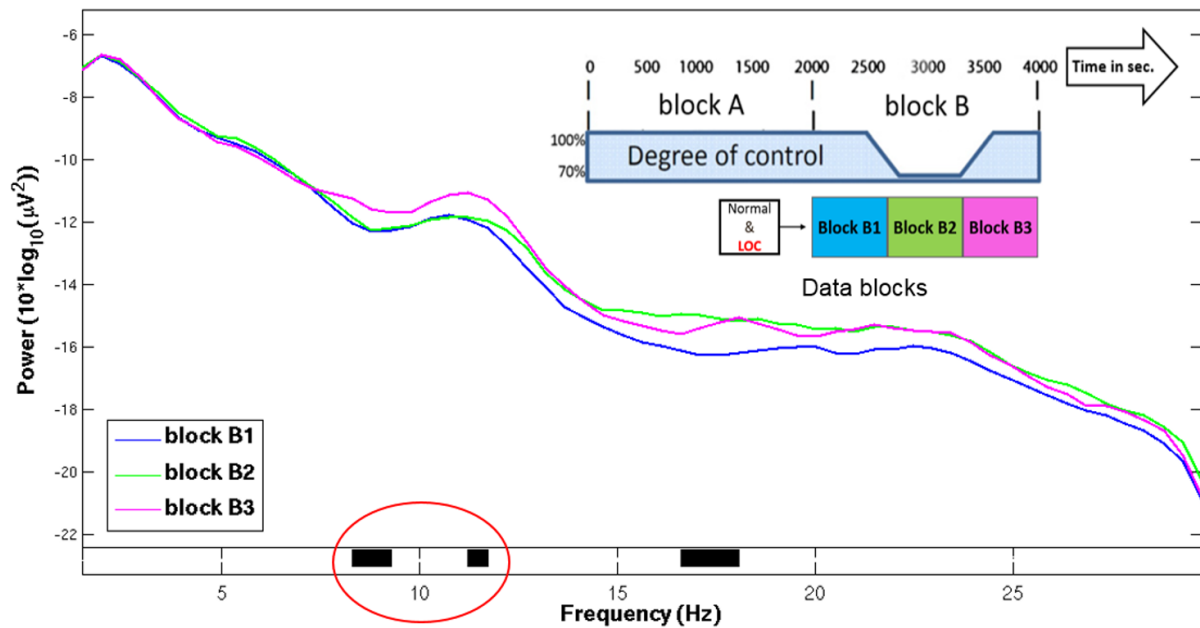
Figure 39. Activity of the right BA6 cluster over the course of the second half of the experiment. I divided the data from the second half of the experiment into three parts (top right) and found that the overall activity (including normal trials and loss of control [LOC] trials) in the cluster exhibited increased alpha activity in the last part of the experiment, which was during a phase where LOC already disappeared.

This finding indicates that, in addition to the affective responses during the LOC phase of the experiment there was another dominance related effect that appeared with some time delay. It appeared when the LOC was already decreasing again and thereby does not represent an emotional response, but more of a mood related phenomenon. This finding will be further discussed in **section 7.3.4**.

6.3.2 Additional Chapter: Examining Cognitive Load during perceived Loss of Control

In cooperation with Dr. Zander and colleagues (Zander et al., in preparation), WML was investigated using the above mentioned dataset. According to established findings in the context of EEG WML detection (Klimesch, 1999; Scharinger, Kammerer, & Gerjets, 2012; Walter et al., 2013), the authors could show that changes in brain activity due to LOC can be attributed to an increase in WML as reflected in increased frontal theta activity as well as decreased parietal alpha activity (see **Figure 40**).
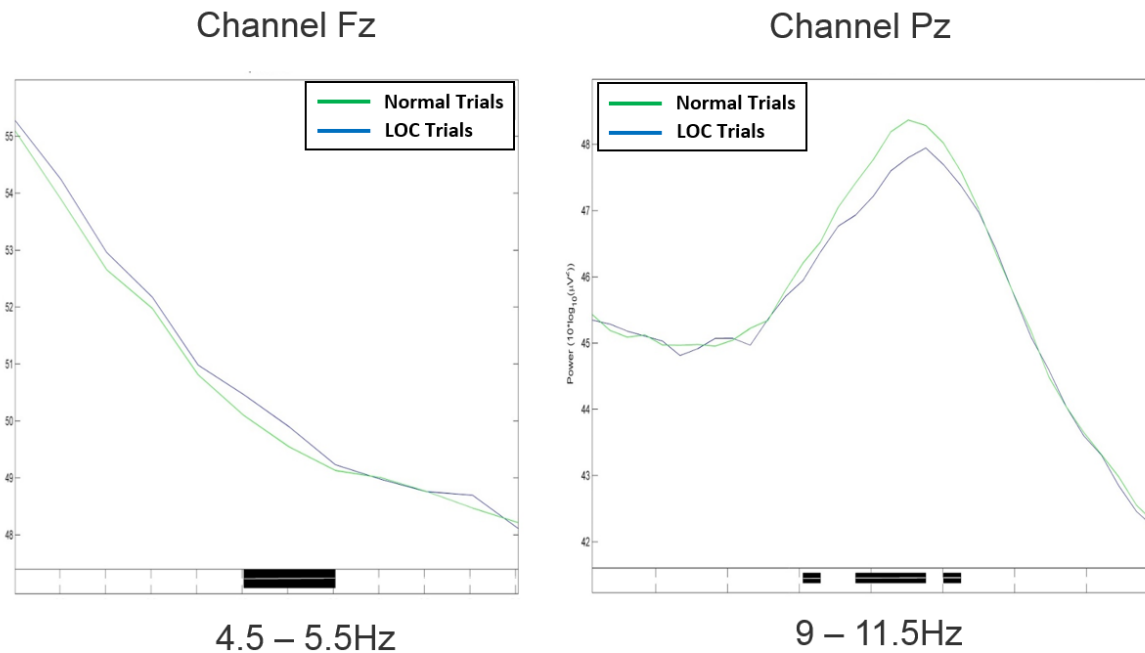
Figure 40. Spectra of channels Fz (left) and Pz (right) computed with EEGLAB's fast Fourier transformation (FFT) and an indicator for statistical significance (p<.05). It can be seen that in the loss of control condition (LOC) theta power on Fz is significantly higher, while alpha activity on Pz is significantly lower than in the condition of full control.

Corresponding electrocortical activity was traced back using ICA and dipole modeling to areas within BA 24 (anterior cingulate cortex or ACC; see **Figure 41**) and BA 31 (posterior cingulate cortex; see **Figure 42**), which are often associated with changes in WML.

Prominently, the ACC has been linked to ERP components reflecting error processing, said to be induced by unexpected action outcomes (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Holroyd & Coles, 2002). Bush, Luu, & Posner (2000) argue that the generation of such error signals is not only a sign of top-down processing, but also of affective mental processing. This is in accordance with the LOC paradigm used here: LOC is induced by (unexpected) machine errors, and it can be expected that these lead to an affective reaction (anger, frustration).

The location of the parietal alpha cluster was reconstructed in or near the border of the posterior cingulate cortex and the cuneus, not far from the intraparietal sulcus. A number of findings relate this area to WML, supporting its identification in this study. In an fMRI study, Leech, Kamourieh, Beckmann, & Sharp (2011) showed that the activity of the posterior cingulate cortex reflects task complexity in workload tasks such

as the N-back task. The cuneus has been shown to have different spectral properties in the alpha band during workload tasks (Michels, Moazami-Goudarzi, Jeanmonod, & Sarnthein, 2008). MacDonald, Cohen, Stenger, & Carter (2000) reasoned that this specific area (BA24) might house a mechanism for performance monitoring during repetitive tasks with a specific regard to incongruence in stimuli.
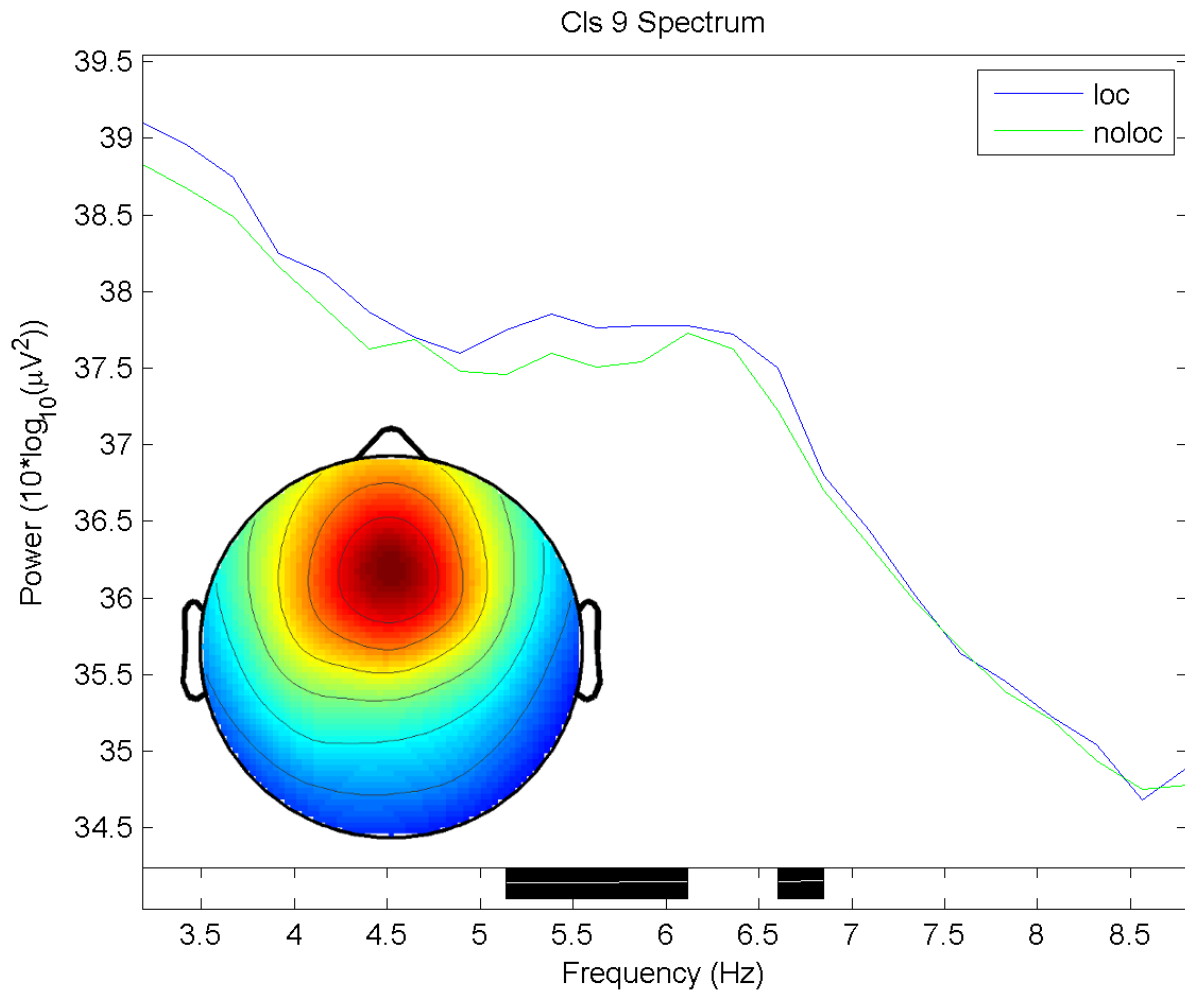
Figure 41. Scalp projection and spectral power of frontal cluster.
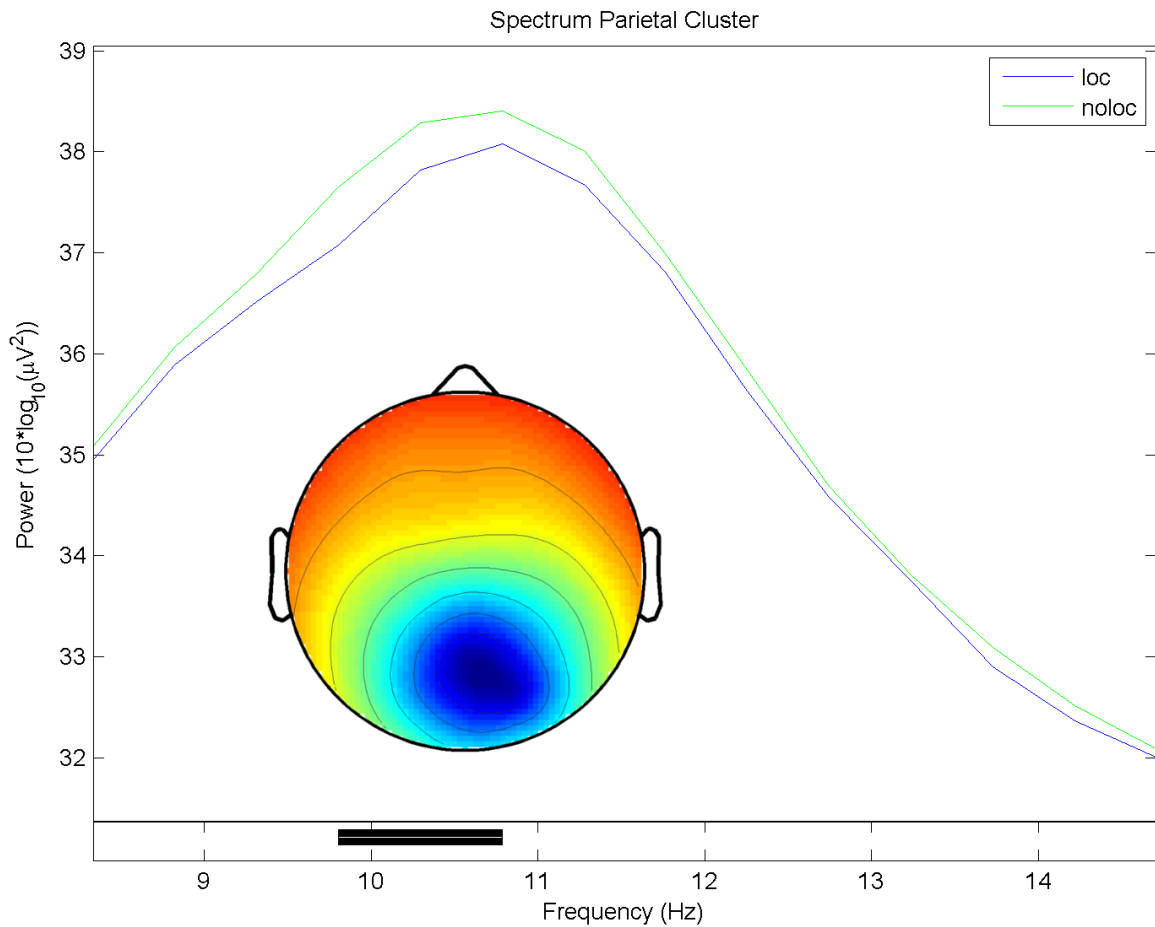
Figure 42. Scalp projection and spectral power of parietal cluster.

A passive BCI discriminated phases with full control from that with perceived LOC, achieving a mean classification accuracy of more than seventy percent (72.3%). The classifier output indicated that WML increased during LOC, but did not return to normal after full control was restored. These findings therefore do not appear to be specific to LOC itself. In the final, full-control phase of the experiment, participants may have been occupied adjusting once again to a changed situation. This indicates that affective responses and their interactions with WML need to be considered as well, which will be part of **section 7.3.4** in the general discussion.

7   General Discussion

This section will start with a discussion of all findings related to WML, affective valence and dominance. Following this, the interactions between these processes will be reviewed. This also contains an attempt to integrate findings from three different studies into the proposed framework. This section will finish with a conclusion that tries to answer all RQs.

7.1   Working Memory Load

At the beginning I will review the results from the subjective and performance measures. After that I will explain the EEG measures for WML like frontal theta activity and parietal alpha activity. Finally, I will reflect on the limitations of these findings and also provide a short outlook for future research.

7.1.1   Subjective and Performance Measures related to Working Memory Load

As expected, I found that increased WML, due to an increase in the load factor of the emoback task, resulted in higher subjective workload ratings. This finding is reassuring, especially since some subjects reported difficulties judging the WML with regard to the difficulty of the emoback task. Some subjects even reported that they experienced the 1-back conditions as more demanding, due to their monotonous nature.

As predicted, my analyses revealed that task performance was impaired due to an increase in WML in the difficult (2-back) conditions. This was reflected in reduced accuracies as well as increased reaction times.

7.1.2   EEG Working Memory Load Measures

Increased WML as induced by the difficult conditions was also reflected in EEG measures used to infer WML. As found in previous studies (Klimesch, 1999; Scharinger et al., 2012), frontal theta activity was increased during high WML (2-back).

Interestingly, the negative affective valence conditions also had an effect on EEG features used to infer WML in the participants. However, contrary to WML, the negative affective valence conditions resulted in decreased frontal theta activity. This finding will be further discussed in **section 7.3.3**.

I will now review the findings for the second workload related EEG measure, parietal alpha activity. As hypothesized, I found that parietal alpha activity was reduced during increased WML. Furthermore, my analyses revealed that negative affective valence also resulted in decreased parietal alpha activity, which will also be further discussed in **section 7.3.3**. In the following section, I will discuss the classification of WML.

7.1.3 Classification of Working Memory Load

The results of the second emoback study demonstrate that WML can be classified under affective valence, providing classification accuracies which can be used in the context of neurotutors. Overall, classification accuracies for WML passed the seventy percent performance threshold, which is widely used for the evaluation of active BCIs (Kübler et al., 2001). This classification performance was already achieved at the single-trial level. Integration over time, using multiple (up to 5) trials to make a given class label estimate, further improved classification performance to almost ninety percent (87.7%). This performance level is comparable to previous research that also used data integration over time to boost performance (see Brouwer et al., 2012). Notably, my results show that using only two trials to integrate data over time led to the most substantial improvements. These two trials represent just five seconds of data, which highlights that our approach is usable in the context of neurotutors, since a robust estimate of the current user state every couple of seconds seems sufficient in such settings.

In line with previous research (see **section 4.3.3.2**), my analyses revealed that frontal theta and parietal alpha features were most important for the discrimination of WML, even under different affective valence. Nevertheless, the relevance of different EEG features strongly depended on the specific affective context. Activity in the frontal theta band appeared most relevant for the discrimination of WML under negative affective valence (see **Figure 27**). This might be explained by the effects found in the

first emoback study, which indicated that negative affective valence led to reduced frontal theta activity (see **section 6.1**). The overall reduced frontal theta activity might have improved the discriminability of the frontal theta features, via a reduction of the variance in this measure. Somewhat similar results have been found in a study presented at the Berlin BCI Workshop 2012 (Grissmann et al., 2012). The authors used motor imagery as well as steady-state-visually evoked potentials (SSVEPs) to train a BCI intended for two-dimensional cursor control and found that classification performance was improved when using the mixed conditions as training data. Their analyses revealed that the signal produced by the SSVEPs did overlap with signals originating from the sensorimotor cortex due to the imagery of motor actions. The authors applied a common spatial pattern filter (Ramoser, Müller-Gerking, & Pfurtscheller, 2000), which could have helped to prune the features in the motor cortex and thereby improved the classification performance.

In contrast to the classification of WML under negative affective valence, WML discrimination under neutral affective valence mostly depended on alpha band activity (see **Figure 26**). The markedly different patterns found for each specific affective context already display the impact of the different affective valence on the features used to infer WML. This was also reflected in the reduced classification performances that were found under positive as well as negative affective valence. These findings are partially comparable to research conducted by Mühl, Jeunet, & Lotte (2014). They used an N-back task to generate WML while social stress was induced using a stress induction protocol based on the Trier Social Stress Task (Kirschbaum et al., 1993). EEG features from the frequency-domain as well as the time-domain have been used to infer WML. The authors concluded that classification methods can be transferred across affective contexts, although only with reduced classification accuracies. In contrast to that, I found that the classifiers used in my study were highly context-specific. The classifiers did not generalize well across different affective contexts. One potential reason for these conflicting results might be the use of additional frequency bands in the study by Mühl, Jeunet, & Lotte (2014). I deliberately avoided the use of frequency bands commonly not used to infer WML to avoid a bias in the results which might occur due to external factors like muscular artifacts. A second reason for the different results could be the type of affect induction used in their paradigm. While I chose to use a task that is inherently activating cognitive as well as

affective processes, the affect induction used by Mühl, Jeunet, & Lotte (2014) was based on a change of the context due to the simulation of an evaluation scenario.

Additionally I found strong inter-individual differences in the recorded EEG data. They highlight the current need for customized systems which are individualized for each single subject and context. The next section will shortly describe the limitations of these findings.

7.1.4 Limitations and Outlook for the Detection of Working Memory Load

The failure to generate WML classifiers that generalize well under different affective contexts raises various concerns, which affect the applicability of such approaches in the context of neurotutors. The creation of training data that is required by machine learning approaches is usually quite tedious. If it turns out that robust detection of WML using the approach described above indeed requires the creation of training data for each specific context, then it seems that we are either forced to focus on very specific application scenarios or develop new approaches that are not so context sensitive. These topics will be further addressed in **section 7**.

7.2 Affective Valence and Dominance

At the beginning I will review the results from the subjective measures concerning the valence and dominance dimension in the emoback studies. After that results from the performance measures in relation to affective valence in the emoback paradigm will be discussed. Following that, I will review the EEG measure for affective valence and dominance, namely the FAA. After that the results from the affective valence classification will be discussed. Finally, this section will describe the limitations concerning these findings and also provide a brief outlook for future studies.

7.2.1 Subjective Measures related to the Valence and the Dominance Dimension

As hypothesized, I also found a significant effect of affective valence on ratings from the valence dimension of the SAM. Notably, the successful induction of positive affective valence shows that the emoback paradigm worked in the intended way. The induction of positive emotions can be very difficult to achieve in laboratory environments, since positive emotions usually arise in a specific context which is difficult to represent in a single picture (Kim & Hamann, 2007).

Unfortunately, I could not acquire reliable subjective dominance ratings during the experiment (see **section 6.1.1**). The subjects were too unfamiliar with the concept of dominance and therefore mostly chose the middle category.

7.2.2 Performance Measures in the Context of Affective Valence

In the emoback studies, I found that negative affective valence resulted in decreased performance, as reflected in the results of both performance measures. These results can be seen as further evidence showing that cognitive and affective components of the proposed framework interact with each other.

This finding is in line with research by Passarotti et al. (2011) who used face stimuli in an affective N-back task. More precisely, the authors found that angry faces used as stimuli resulted in increased reaction times. Such results could be explained through the concept of hot cognition which states that decision making can be impaired by stimuli with strong emotional meaning or content (Huijbregts, Warren, De Sonneville, and Swaab-Barneveld, 2008). Thus, this framework assumes that the processing of stimuli with strong positive or negative affective valence can divert limited cognitive resources from the primary task and thereby decrease performance.

One example for such a diversion of cognitive resources would be the process of the automatic capturing of attention, which has been found in a study by Levens and Gotlib (2010). The authors found that strongly valenced stimuli have a tendency to stay longer active in WM. This might potentially interfere with different EFs required by the emoback paradigm (Miyake et al., 2000), mainly with the core EFs of updating and inhibition. The emoback paradigm demands that participants constantly refresh

(update) the content in WM by adding new stimuli and replacing (inhibit) old ones. Furthermore, participants were also required to switch (shift) between a simple identity-matching task (before reacting to the stimulus) and an updating task (after reacting to the stimulus). Strongly negative valenced stimuli could have caught the attention of the subjects, diverting it to the negative valenced stimuli and away from the primary task. This could have slowed down the rearrangement of stimuli in WM, thereby resulting in impaired task performance.

7.2.3 EEG Measure for Affective Valence – Frontal Alpha Asymmetry

The results from the subjective measures as well as the performance measures clearly demonstrate that the paradigm worked in the intended way. I was able to successfully induce WML and affective valence. Nevertheless, I could not find any significant effect using the FAA measure to detect affective valence during the emoback task. Despite a large body of evidence supporting this phenomenon, this study does not represent the first incidence where researchers unsuccessfully attempted to use this measure to infer changes in affective valence. Various studies found that the FAA is not always observable in response to affective stimulation (Fairclough & Roberts, 2011; Kop et al., 2011; Sarlo, Buodo, Poli, & Palomba, 2005; Schutter, Putman, Hermans, & Van Honk, 2001; Winkler, Jäger, Mihajlovic, & Tsoneva, 2010).

For instance, recent research conducted by Hettich, Bolinger, Matuz, Birbaumer, and Spüler (2016) used auditory stimuli from the international affective digitized sounds database (IADS-2; Bradley & Lang, 2007) to induce affective valence. They used the LPP and the FAA as indicators for affective valence and found that only the LPP produced brain responses that could be used to automatically detect the affective valence during human-machine interaction. Their conclusion is similar to mine, stating that EEG measures used to infer the ongoing user state are highly context-specific and therefore depend on a given experimental paradigm. This leads to another potential cause for the non-responsiveness of the FAA under WML. Brain laterality responses have been measured during rest or passive viewing in most previous studies (Ahern & Schwartz, 1985; Davidson & Wheeler, 1992; Huang et al., 2012; Lin, Wang, Wu, Jeng, and Chen, 2009; Ramirez & Vamvakousis, 2012). The paradigm used in my studies tried to infer affective valence using the FAA while

participants were simultaneously engaged in a WML task. This could have created interference, especially since the FAA is known to be influenced by WML (Briesemeister, Tamm, Heine, and Jacobs, 2013). For instance, a study by Baldwin and Penaranda (2012) found that increased task difficulty resulted in increased left frontal activity. My analyses revealed that performance was impaired by negative affective valence. However, this impairment should have resulted in increased right frontal activity, since right frontal activity is related to the experience of negative affective valence. I therefore assume that it might be possible that both influences, WML and negative affective valence, influenced the same FAA measure. However, these influences could have pointed in opposing directions, thereby potentially canceling each other out and masking any effects.

This finding further highlight the need for frameworks that enable us to account for such specific contexts. Additionally, it underscores the necessity for a re-evaluation of results from studies which were highly controlled and therefore did not account for the potential influence of such context factors. Only with the help of such studies can we discover which results can be used in more natural environments outside the lab.

7.2.4 Frontal Alpha Asymmetry as Indicator of Affective Dominance Responses

The analyses of the LOC study revealed that trials with incorrect stimulus rotation (LOC) were accompanied with increased left frontal activity due to increased right frontal alpha activity. Left frontal activity is usually associated with increased affective dominance as well as positive affective valence. While the shift towards positive affective valence seems puzzling at first, the shift toward more dominance could be explained with feelings of anger (Demaree et al., 2005), which might have been experienced by participants during LOC. Such approach-related responses appear to reflect engagement (Harmon-Jones et al., 2010), which could indicate that subjects made efforts to regain control of the system after LOC. As previously learned strategies that were used to solve the task did not work anymore, it is highly likely that anger arose that fueled the desire to regain control. **Figure 43** helps to better understand the typical learning trajectory over the course of the experiment.
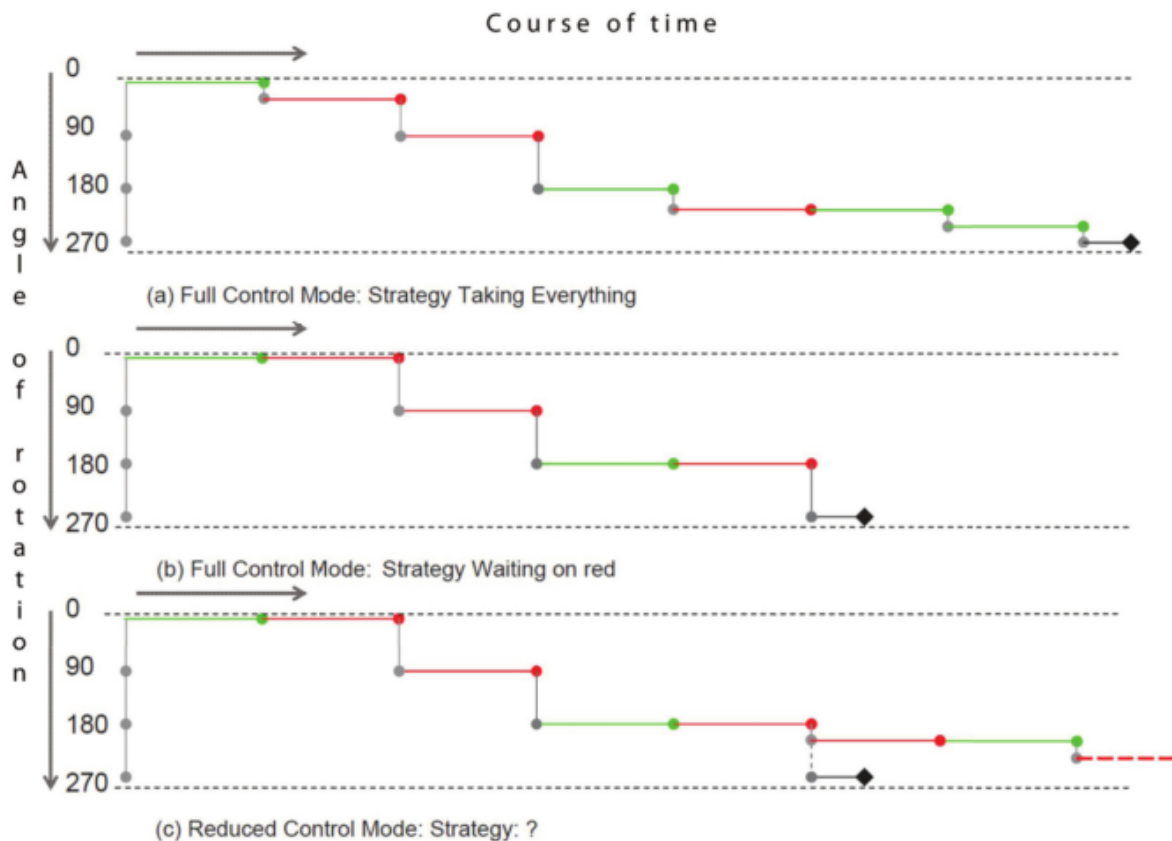
Figure 43. Strategies developed during the LOC paradigm. Time is displayed horizontally and the vertical axis shows the angle of rotation (color coded). At the top the full control mode and the 'taking everything' strategy is displayed. The middle shows the more efficient 'waiting on red' strategy in the full control mode. At the bottom the potential behavior of somebody who tries to implement the 'waiting on red' strategy during LOC is displayed. However, as this example illustrates, this strategy does not work during LOC, requiring the user to rethink his strategy accordingly (Zander, 2011).

At the beginning of the experiment, most subjects took all presented rotation opportunities. After a while the participants realized that they can reach the goal faster if they waited for red, which indicated a larger rotation angle. However, during LOC the subjects could not apply the learned strategies anymore, since the system did not respond in the expected manner. This seems to have forced the subjects to either try to regain control or withdraw fully from the task instead. Additionally, related analyses from Zander et al. (in preparation; see **section 6.3.2**) further support this assumption, since they found that WML was increased during LOC. The seemingly erroneous behavior of the system did not appear to result in task disengagement, since task disengagement should have been reflected in decreased left frontal activation as well

as reduced WML. This finding further supports the assumption of additional efforts on behalf of the subjects that usually accompany high dominance behavior.

Nonetheless, the discovered laterality effect could also be interpreted as shift toward positive valence, which appears to be counter-intuitive, since approach-related emotions due to LOC should have been accompanied by a shift toward negative affective valence. Results from the source localization analysis will help to resolve this apparent inconsistency.

7.2.5 Source Localization Analysis of the LOC Data

Analyses of the component cluster activities revealed that the laterality effect found in the alpha power difference plots (see **Figure 34**; right) most closely resembled the scalp projection of a component cluster that was located in or near the right BA6 (see **Figure 37**). Interestingly, this BA6 cluster showed decreased alpha activity in the lower alpha band and increased activity in the upper alpha band during LOC. This differential effect supports the assumption that the potential experience of anger during LOC trials was likely accompanied by emotional responses with high dominance and negative valence. This interpretation is in line with Reuderink, Mühl, and Poel (2013) who found similar results in related research that also manipulated the interaction in a human-machine system. The authors measured valence and dominance during game play and found that activity in the lower alpha band was related to affective valence and activity in the upper alpha band was related to affective dominance. Such findings suggest that the FAA response is more complex than previously assumed. A separate analysis of different alpha bands in future studies might help to disentangle different responses that are usually associated with the FAA measure.

As indicated by previous research, emotions are often related to certain motor response tendencies (Önal-Hartmann et al., 2012). Humans generally push away negative stimuli and reach for positive stimuli. One region of the brain which is assumed to play a major role in the planning of such motor responses is BA6 which includes the premotor cortex as well as the SMA. The centroid of the right BA6 cluster was located closest to the right SMA (see red dot in individual dipole locations in **Figure 37**; top right). In past research, the SMA has been shown to be active in response to unclear stimuli (Watson et al., 2013). I assume that due to the random nature of the

LOC paradigm the interpretation of the stimuli color was affected during LOC, leading to an uncertainty about which response could be expected from the system. Furthermore, SMA activity was also found during unimanual and stimulus-cued movements (Picard & Strick, 2003). Such response behavior is also part of the LOC paradigm, which required participants to react via unimanual key press to letter stimuli. Additionally, a study by Oliveri et al. (2003) revealed that the use of emotional stimuli resulted in enhanced motor evoked potentials in the SMA. Since I assume that the LOC was accompanied by affective reactions, one can expect that the LOC stimuli did function as emotion evoking stimuli. This could further explain the increased activity in the right BA6 cluster. Finally, the SMA has been shown to exhibit increased activity during emotional conflict, which was found in a study using an affective flanker task (Ochsner, Hughes, Robertson, Cooper, & Gabrieli, 2009). The random nature of the occurring experience of LOC makes it highly plausible to assume that some kind of emotional conflict was induced via the LOC paradigm, further indicating that this component cluster was the main contributor to the laterality effect I found at the electrode level. Additional support for my interpretations is provided by the results of the analysis of the homologous BA6 component cluster located on the left hemisphere. This component cluster did not exhibit increased alpha band activity, which suggests that it did not contribute to the laterality effect found at the electrode level. Furthermore, correlational analyses revealed that the scalp projection of this component cluster did not show a strong similarity with the difference plot (r=0.17, p <.001, sig.; see **Table 1**, **Figure 38** and **Figure 34**; right), which is additional evidence for a main contribution of the right BA6 cluster to the found laterality effect. The next section will discuss the findings from the affective valence classification from the second emoback study.

7.2.6 Classification of Affective Valence

In the first emoback study I only analyzed effects on the group level and the large variance in EEG measures hinted at substantial inter-individual differences in the recorded data sample (see **Figure 19**). Accordingly, I cannot rule out laterality effects that existed on an individual level. Classification approaches commonly used in the field of BCIs usually operate on a single subject level (Allison et al., 2011; Hwang, Kim, Choi, & Im, 2013; Leeb et al., 2007) and therefore might help reveal individual

responses, which cannot be found at the group level. This analysis of individual EEG responses was part of the second emoback study (see **section 6.2**), which will be discussed in this section.

My analyses revealed that affective valence can be classified under WML, resulting in classification accuracies which indicate that our approach is feasible in the context of neurotutors. Mean classification accuracies were above the seventy percent threshold (Kübler et al., 2001) for single-trial classification using the normal approach that used global alpha power changes as features. Performance of the normal approach was further improved via data integration over time, resulting in peak accuracies of close to ninety percent (89.4%). To further investigate the potential use of FAAs in the context of neurotutors, I created additional features based on difference scores that resemble the FAA measure (see **section 6.2.1**). Using the FAA approach resulted in overall classification results that just failed to exceed the seventy percent threshold at the single-trial level under low WML (68.4%). However, data integration over time helped to improve performance in this approach as well, reaching accuracies as high as eighty percent (80.5%). The achieved performance level is comparable to previous research that tried to classify pure affective states (Kim et al., 2013). Similar to the classification of WML, data integration over time using two trials provided the largest improvements for both valence classification approaches.

Surprisingly, I found that parietal alpha activity was most relevant for the discrimination of affective valence in the normal approach (see **Figure 29** and **Figure 30**). This finding seems to indicate that the normal approach was most sensitive toward changes in WL features which reacted to the specific affective context, something that could prove highly problematic if one wants to differentiate between WML and affective valence, since both seem to modulate the same EEG response. Such findings highlight the necessity to limit data input for the classifier, which was the intention behind the FAA approach. Surprisingly, I found that discrimination of affective valence in the FAA approach was mostly due to frontal alpha activity in the upper alpha band (see **Figure 31** and **Figure 32**). Reuderink, Mühl, and Poel (2013) also used laterality measures in a related investigation that also manipulated control over a human-machine interaction system. They used a variant of the popular Pacman game and induced frustration by ignoring a certain amount of keyboard input. With the use of the SAM scales (Bradley

& Lang, 1994) they found that activity in the lower alpha band was related to the valence dimension, while activity in the upper alpha band was related to the dominance dimension. Similar results have also been found in my third study that also manipulated the level of control during human-machine interaction (LOC study; see **section 0**). The subjects in the emoback study primarily reported feelings of disgust in response to the negative valenced stimuli (e.g., mutilations). Since there was no way to avoid exposure with the stimuli without stopping the experiment, one could assume that the experience of negative valence was accompanied by feelings of low dominance. Nevertheless, since I could not acquire reliable subjective dominance ratings during the experiment (see **section 6.1.1**), I can only make assumptions about potential causes for these results. Furthermore, one has to keep in mind that, to the best of my knowledge, only one study (besides study 3 from this thesis) simultaneously investigated valence and dominance using FAAs. It is therefore difficult to draw any definite conclusions without further research.

Interestingly, I found that affective valence classification was improved under high WML for both approaches. This finding might result from the monotonous nature of the task in the low WML conditions. Subjects reported the 1-back conditions as very boring, which could have decreased the level of emotional reactivity, thereby interfering with the intended emotion induction. Nevertheless, this assumption is not backed up by the subjective ratings, since I did not find an effect of WML on affective valence ratings. Therefore, this assumption would need further verification as well.

Cross-task classification revealed that my classifiers were task specific and did not respond to changes in WML that did not involve changes in affective valence. Furthermore, I found that the classifiers did generalize well across different cognitive contexts. However, I do not want to draw premature conclusions since I found that the affective valence conditions were confounded with time on a single-subject level, which will be discussed in the next section.

7.2.7 Limitations and Outlook for the Detection of Affective Valence and Dominance

Based on the reasoning outlined above, I assume that the reason for the inability to find effects at the group level in the first emoback study using the FAA measure was

the concurrent induction of WML in our paradigm. To test this assumption one could introduce a passive viewing condition into the paradigm I developed. If such an experiment would find that the FAA measure can be used during phases without additional cognitive processing, one could account for this in the design of a human-machine system like a neurotutor. This could be done, for example, by introducing short breaks between tasks that could then be used to infer the current affective user state.

Unfortunately, the affective valence conditions in the emoback paradigm turned out to be confounded with time-on-task on a single-subject level. The use of the FAA approach should have eliminated most of this influence, because the FAA only measures differences between the left and the right hemisphere and time-on-task effects are assumed to occur at a global level. Nonetheless, since it is impossible to fully disentangle effects which are due to the affective induction from those which might occurred due to the time-on-task, I abstain from further interpretations of the findings from the affective valence classification.

I chose to induce emotions block-wise in my paradigm because it can be difficult to induce strong affective reactions repeatedly without risking habituation toward the emotion induction. Nevertheless, I suggest that future research should use short blocks (up to 10 stimuli) of trials from a specific combination of N-back levels and affective valence instead of the block design used in this study. Such an approach would allow randomization of all conditions, thereby eliminating the time-confounds present in this dataset.

7.3 Interactions between Working Memory Load and Affective Processes related to the Valence and Dominance Dimension

While the previous sections have already shown that WML and affective valence do interact with each other, this section will further elaborate on those findings. At the beginning I will review the results from the subjective measures. After that performance measures will be discussed. Then I will discuss the findings from the EEG measures. Finally, I will attempt to use the proposed framework to integrate the findings from three studies that all analyzed the same LOC dataset.

7.3.1 Subjective Ratings indicating Interactions between Working Memory Load and Affective Valence

There was a significant effect of negative affective valence on WML ratings, indicating that negative affective valence led to the subjective experience of increased WML. This can be seen as further evidence that indicates that WML can be influenced by affective valence and that this influence was strong enough to be subjectively perceivable by the subjects. However, based on such subjective measures alone, one does not know much about the potential causes of such interactions.

7.3.2 Performance Measures in the Context of an Interaction between Working Memory Load and Affective Valence

My analyses revealed a significant interaction effect for accuracy between WML and affective valence. I found that increased WML resulted in decreased accuracies (d=0.50), especially during negative affective valence (d=0.69). Our results are in line with previous findings from a study by Kopf, Dresler, Reicherts, Herrmann, and Reif (2013) who used affective words as stimuli in a N-back task. They also found decreased accuracies in the difficult conditions during negative affective valence. This is also in line with a study by MacNamara, Ferri, and Hajcak (2011) that used emotional pictures as distractors in an emoback task. The authors also found that the emotional content of negative valenced stimuli further aggravated the negative impact of WML on performance. Such results seem to demonstrate that WML as well as negative affective valence compete for limited cognitive resources. It appears that such an increased drain from a limited pool of mental resources can result in decreased performance. These findings could also be explained using the capacity model from Ellis and Ashbrook (1989). It seems that negative affective valence influenced the allocation of available attentional resources toward the task and thereby impaired task performance. However, I only found a small (η^2=0.14) interaction effect for accuracies, thus the available evidence for this interpretation is rather limited. Additionally, as was the case with the results from the subjective ratings, it is difficult to draw any strong conclusions based on performance measures alone. Based on such measures one can only make vague assumptions on how negative affective valence might have

impaired task performance. I will now turn to the results from the EEG analysis, which will help to develop a more conclusive picture.

7.3.3 EEG Measures in the Context of an Interaction between Working Memory Load and Affective Valence

The analysis of the emoback studies have shown that both WML and affective valence interact with each other. However, these interactions appear quite differently for both components of the proposed framework. While the detection of WML was limited under different affective valence, the detection of affective valence under WML was not possible at the group level using established EEG measures.

Furthermore, negative affective valence had an influence on measures used to infer WML. Negative affective valence resulted in decreased frontal theta activity as well as decreased parietal alpha activity. If negative affective valence would impair performance via the production of additional WML, one would expect to find increased frontal theta activity during negative affective valence. However, the decreased frontal theta activity under negative affective valence indicates that the detrimental effect on performance likely has different reasons than the performance decrease under high WML. Accordingly, I assume that the negative affective valenced stimuli in the emoback study interfered with task processing through a reduction of activity in the frontal attentional control network.

The findings of the meta-analysis conducted by Krain, Wilson, Arbuckle, Castellanos, and Milhama (2006) could provide an explanation for this frontal theta effect. The authors found that the orbitofrontal cortex seems more involved in hot cognitions, while the dorsolateral prefrontal cortex appears to be more relevant for cool cognitions. The orbitofrontal cortex is located directly behind the eyes, which makes it very difficult to detect brain activity that originates from this brain area with the EEG. Accordingly, one could assume that EEG instead mostly captures activity from the dorsolateral prefrontal cortex. As a result, the reduced frontal theta power for negative affective valence found in my study might reflect reduced activity of the dorsolateral prefrontal cortex. Nevertheless, since both regions are in close vicinity and no source localization analysis has been conducted yet, this assumption needs to be tested through additional research.

A different explanation for the reduced activity of the frontal attention network could be the reduced allocation of attentional resources toward the criterion task, as implied by the capacity model from Ellis and Ashbrook (1989). Negative emotions can induce negative thoughts and thereby reduce resources available for task performance (Ellis & Ashbrook, 1988). Frontal brain regions like the ACC are assumed to be involved in EFs (Corbetta & Shulman, 2002; Weissman, Gopalakrishnan, Hazlett, & Woldorff, 2005), which is also reflected in the increased frontal theta activity which was found during LOC (see **section 6.3.2**). Additionally, this part of the brain also seems to interact with other brain structures which are related to affective processing (Bush et al., 2000). According to Pessoa (2009) the resulting brain network integrates control signals with affective information, which can either boost or, as in our case, impair performance.

Finally, this interference could also be explained by the so-called attentional capture effect as found in a study by Hodsoll, Viding, and Lavie (2011). The authors used irrelevant emotional distractor faces in a search task and found that distractor faces slowed down reaction times. Furthermore, previous research has also found that affective stimuli usually carry more relevant information and are therefore more salient (Carretié, 2014), which could be one of the causes for such an attentional capture effect.

Parietal alpha activity was also influenced by negative affective valence, which seems to indicate that negative valenced stimuli can lead to increased processing in parietal storage areas. Similar results have been found in a study by Rämä et al. (2001) who used affective voices in an N-back task. The authors concluded that the parietal cortex is involved in the active maintenance of emotional material. Accordingly, I assume that parietal alpha activity was reduced due to the increased processing of the negative valenced picture stimulus material.

These findings show that studies that only focus on single components of the proposed framework are prone to produce misleading results. The next section will try to provide an example case for the potential benefits of a framework that includes multiple components of an ongoing user state.

7.3.4 Integration of Findings into the proposed Framework

I will now integrate the findings from three studies that analyzed the same LOC dataset into the proposed framework by considering the following three steps: As first step I want to integrate the findings from Zander and Jatzev (2012) into the perception and action component. As second step I will integrate the findings from Zander et al. (in preparation) into the cognitive processes component. As third and final step I will integrate my LOC findings into the affective processes component of the proposed framework.

Analyzing the same dataset, Zander and Jatzev (2012) found a shift in the feature space that was used to detect the motor response that was associated with input during the LOC paradigm. These non-stationarities in the feature space were measured using the Kullback-Leibler divergence (KLD; see **Figure 44**; left), a measure for the difference between two distributions. Furthermore, the non-stationarities that were induced by the LOC were accompanied by decreased pseudo online classification (POC) accuracies of the motor execution response (keyboard press; see **Figure 44**; right).
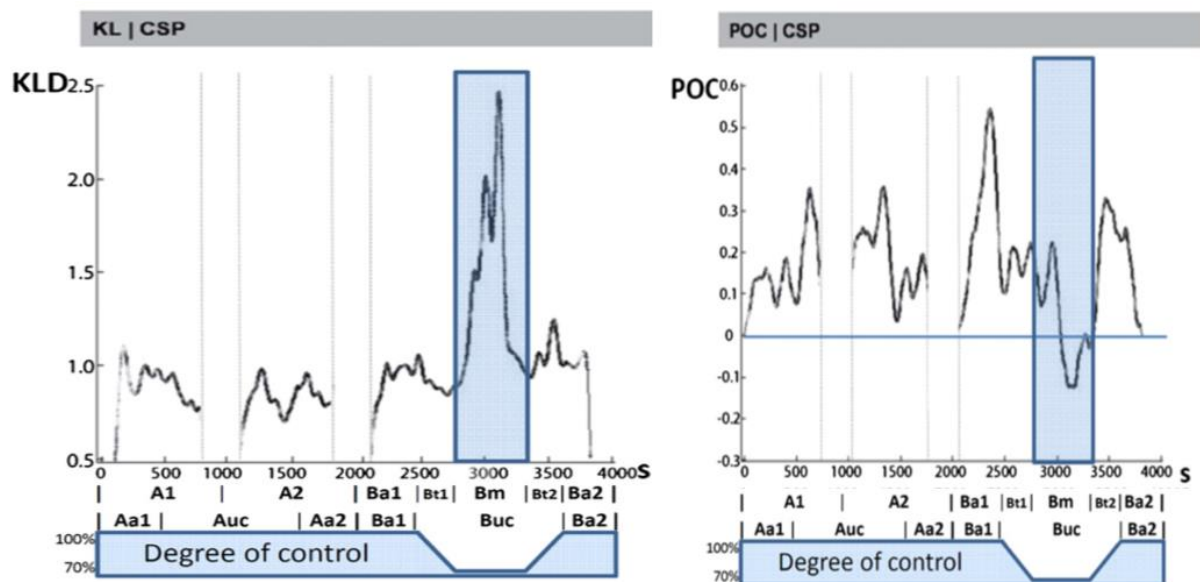
Figure 44. Kullback-Leibler divergence (KLD; left) and pseudo online classification (POC; right) during LOC (modified from Zander & Jatzev, 2012)**.**

Accordingly, the authors concluded that LOC induced non-stationarities in the feature space which then resulted in decreased classification performance. These input related processes are part of the perception and action component of the proposed framework (see **Figure 45**, bottom left).
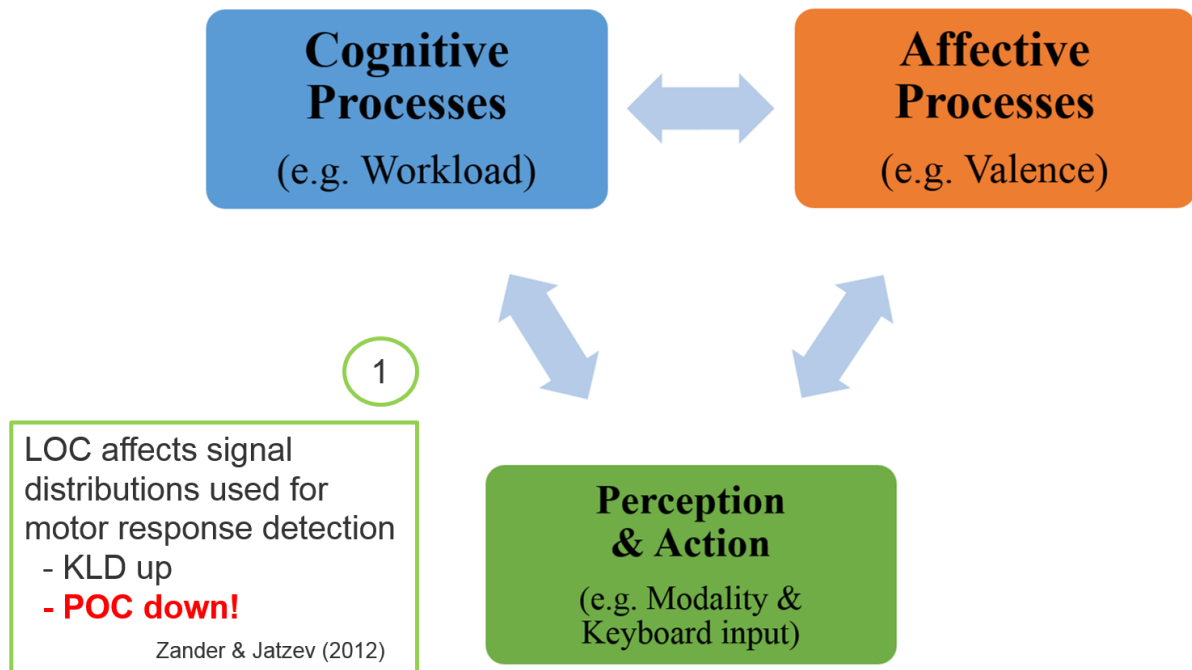

Figure 45. LOC results from Zander and Jatzev (2012) in the context of the proposed framework. The analyses from Zander and Jatzev (2012) focused on the effect of LOC on features used for the detection of user input (keyboard). Since, these effects relate to the interaction with the system, they are part of the perception and action component of the proposed framework (KLD = Kullback-Leibler divergence; POC = pseudo online classification).

Next I want to integrate the results from Zander et al. (in preparation) into the proposed framework. The authors also re-analyzed the same dataset and discovered WML related effects (see **section 6.3.2**), which are part of the cognitive processes component of the proposed framework (see **Figure 47**, left). The authors found increased frontal theta activity (see **Figure 40**; left) as well as reduced parietal alpha activity (see **Figure 40**; right) during LOC. Additionally, the authors found a component cluster with decreased alpha activity located in or near Brodmann Area 31 that projected to areas used for detecting motor responses, contributing to the effect found in Zander & Jatzev (2012; see **Figure 48**; left).
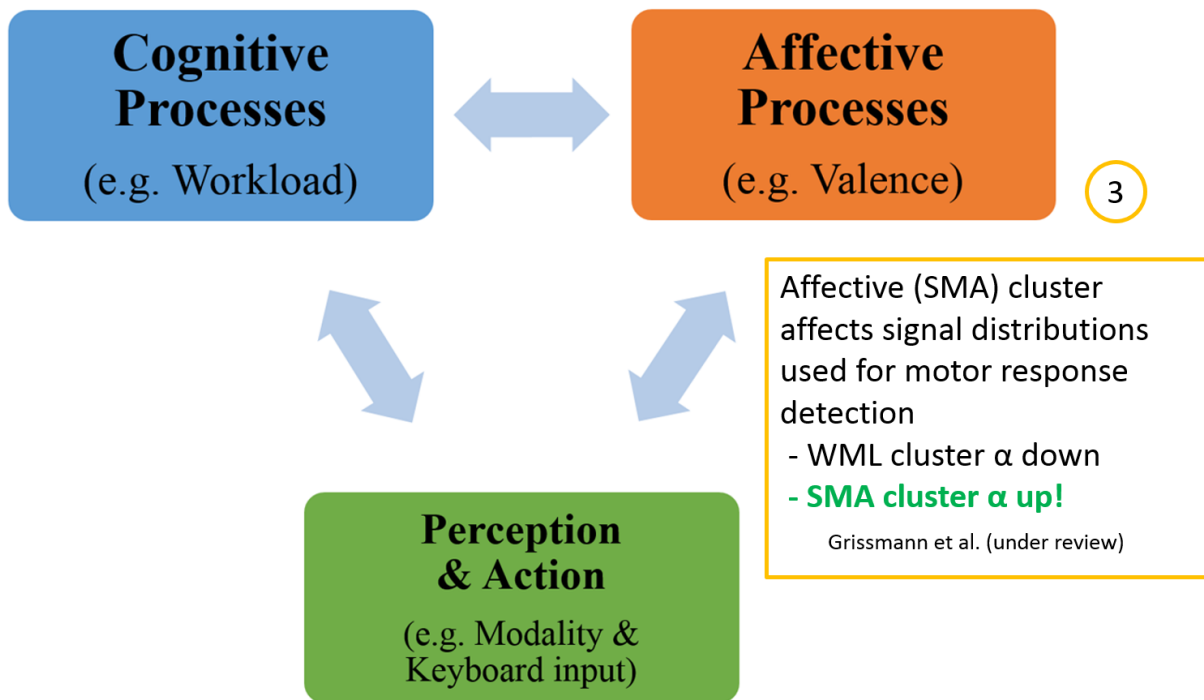
Working Memory Load
Cluster (BA31)
α-down during LOC

Working Memory Load
Classifier
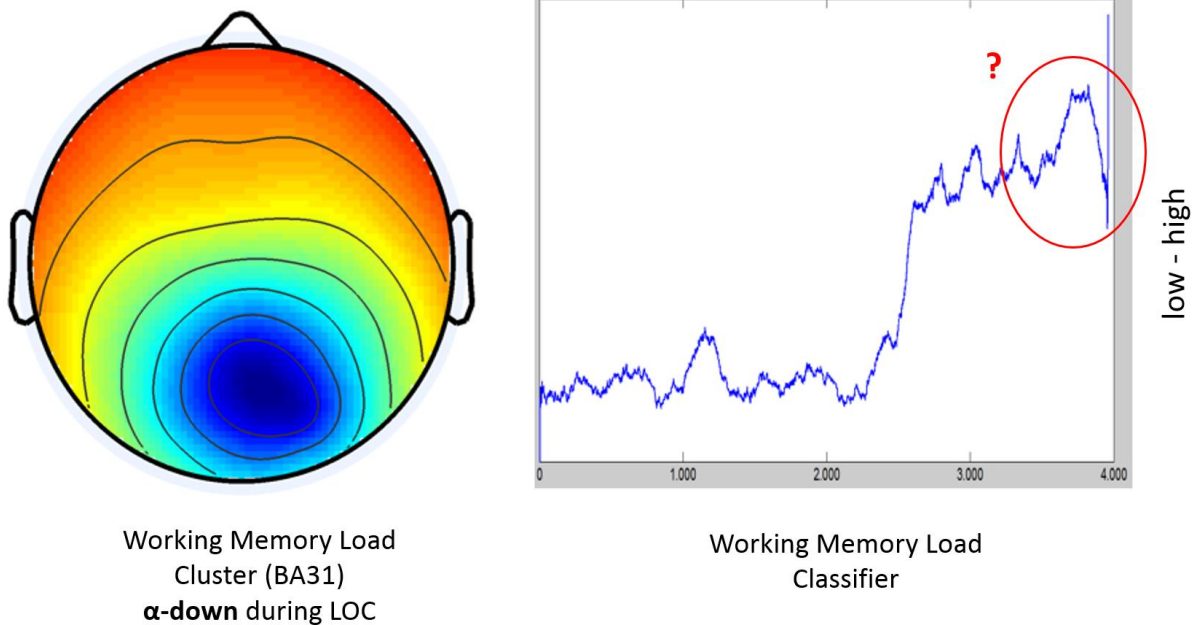
Figure 46. Scalp projection of the parietal alpha cluster during LOC (left) and performance of the working memory load classifier over the course of the experiment (right) (modified from Zander et al., in preparation).

Furthermore, the authors were able to classify the increased WML during LOC by means of a passive BCI. The classifier showed a clear increase of WML with the introduction of LOC, but it exhibited high levels of WML even after LOC (see **Figure 46**; right). This result suggests that the classifier still shows increased WML, even after the signal distributions investigated in Zander and Jatzev (2012) went back to their original state (no LOC). Therefore, the authors concluded that additional processes must have played a role for the non-stationarities found in the signal distributions used for the motor classifier.
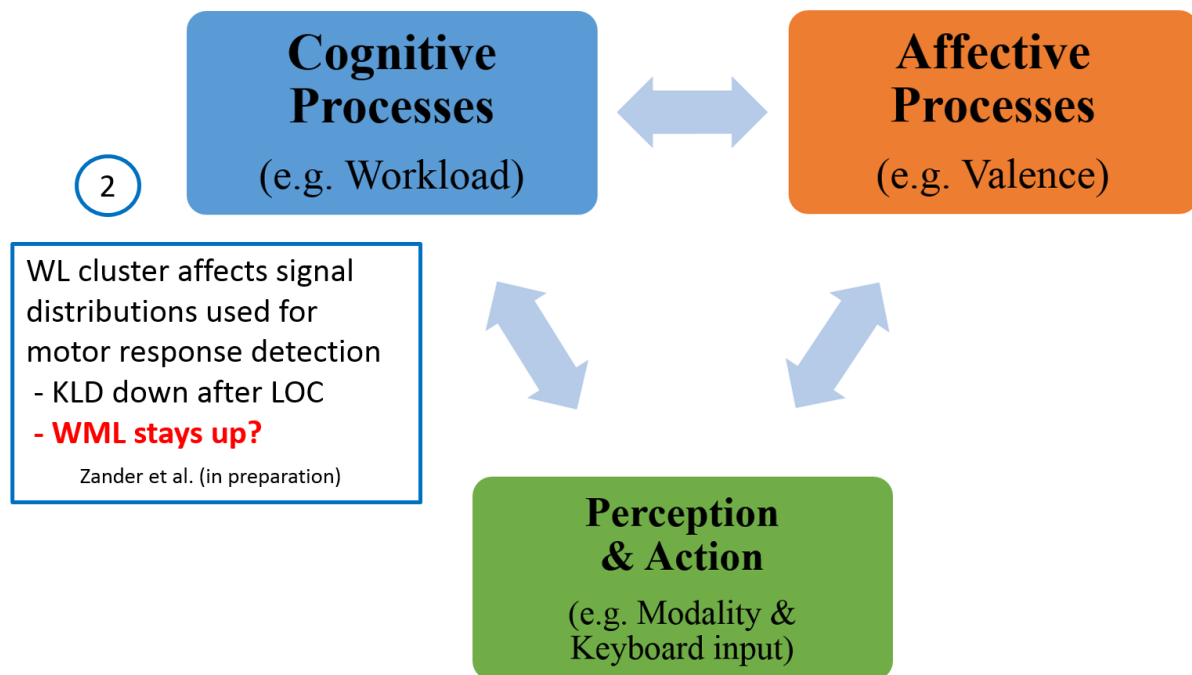
Figure 47. LOC results from Zander et al. (in preparation) in the context of the proposed framework (KLD = Kullback-Leibler divergence; WML = working memory load)**.**

Finally I will make use of the results from the additional section (**6.3.1**) of the LOC study (Grissmann et al., under review) to try to explain why the workload classifier still indicated higher levels of workload at the end of the experiment, but the classification of the motor execution response worked properly again. This is part of the affective processes component of the proposed framework. When inspecting the component cluster that was located in or near the right BA6 (see **Figure 37**) I found that the scalp projection of this cluster also overlapped with scalp areas used for detection of the motor response. However, activity found in this cluster showed increased upper alpha power during and after LOC, while activity in the WML cluster showed reduced alpha power. Thus, it seems likely that the activity in the WML cluster and the activity in the BA6 cluster counter-balanced each other, resulting in a recovery of the KLD as well as the POC. Therefore I conclude that my findings help explain why the KLD and the POC recover after LOC, while the decreased alpha activity induced by increased WML is still present (see **Figure 48**; right).

Figure 48. Results from Grissmann et al. (under review) in the context of the proposed framework (SMA = supplementary motor area; WML = working memory load)**.**

Similar interactions have been found between WML and fatigue. Increased WML resulted in decreased alpha band power, while fatigue led to increases in alpha band power, thereby interfering with WML detection (Roy et al., 2013).

I hereby presented empirical evidence for the necessity to include multiple components into a framework if one wants to track ongoing changes in the brain during natural human-machine interaction. However, additional research is needed to further substantiate these claims. The following section will integrate all of my findings and thereby illustrate how my findings provided answers to the research questions.

7.4   Conclusion

In this dissertation I wanted to investigate if the EEG can be used to infer complex user states that are assumed to arise during natural human-machine interaction. With this aim in mind I investigated if basic components of such complex user states can be separated using measures commonly used in EEG studies. Furthermore, I wanted to test, if said measures are robust or context-specific due to interactions between different basic components of a complex user state.

In the studies of my dissertation, I was able to use established EEG measures to separate cognitive processes from affective processes of the proposed framework in paradigms that evoke complex user states (RQ1). More specifically, I was able to detect WML and affective valence in a controlled study that used an N-back task with affective pictures as stimuli. However, only EEG measures commonly used to infer WML were sensitive to changes in WML as well as affective valence. Additionally, I was able to replicate my findings while inducing complex user states in a less controlled study that used a simplified learning scenario as paradigm. To be more specific, I was able to detect EEG responses related to affective dominance during LOC. In addition to this, I was able to use EEG-based source localization methods to identify cortical sources that were related to the affective dominance response (RQ1a). Analyses conducted in cooperation with colleagues identified WML-related responses in the same dataset and were also able to identify cortical sources associated with these responses (RQ1a). Furthermore, the source localization approach enabled me to separate affective processes related to the valence dimension from affective processes related to the dominance dimension (RQ1b). This helped to reveal effects that were not visible at the sensor level, thereby indicating that the measures used to infer affective responses might exhibit differential effects, which are commonly not accounted for.

Furthermore, I found that EEG measures used to infer certain user states were sensitive to specific context (RQ2). More precisely, I found that EEG measures used to infer WML were sensitive to affective valence, while EEG measures used to infer affective valence were not sensitive to changes in affective valence when WML was imposed simultaneously. With the use of the proposed framework I was able to integrate findings from multiple studies that analyzed the same dataset, thereby demonstrating that the use of such frameworks can help to identify context-sensitivity in EEG measures which is usually not accounted for. More specifically, I was able to integrate findings from three different studies that all analyzed the same LOC dataset. This helped to explain interactions between different processes that are active during the experience of a complex user state. This could not have been achieved without the proposed framework as studies in this field usually only focus on single, narrow user states, thereby neglecting complexity that arises through the interaction of different processes that arise in natural environments.

My analyses revealed that it is possible to automatically detect WML and affective valence with accuracies significantly higher than chance using paradigms that evoke complex user states (RQ3). In detail, I was able to classify WML with accuracies greater than seventy percent at the single-trial level, which was further improved with data integration over time. However, the classification approaches failed to generalize across different affective contexts, indicating the necessity to train specific approaches for each setting. Furthermore, I was able to classify affective valence, achieving accuracies greater than seventy percent at the single-trial level for all but one approach. As was the case with classification of WML, data integration over time did further improve classification performance. Additionally, my analyses revealed that measures used to infer affective valence, which did not exhibit effects at the group level, did show effects at the individual level. However, due to time-confounds in the recorded data, I cannot be certain concerning the validity of these results.

Despite all these interesting findings, my research should only be seen as a first step into a new direction that tries to investigate more realistic human-machine interaction. Before we can fully incorporate such technologies into our daily lives, many important issues still need to be resolved. The next section will discuss some of these issues.

8   Next important Steps towards a reliable Neurotutor

In this section I will discuss crucial steps that need to be taken to develop robust neurotutors that can be used in everyday classrooms. This will be done in seven segments. Firstly, I will discuss the necessity to validate and (hopefully) replicate the findings of this thesis. Secondly, I will shortly describe alternative analytical approaches that might reveal additional aspects in the EEG response. Thirdly, I will point to the necessity to investigate features which are not as established as the ones I used in this thesis. Fourthly, I will shortly illustrate the potential of adding peripheral signals to EEG recordings. Fifthly, I will discuss the need for new, more accessible sensor technology. Sixthly, I will describe potential extensions of the proposed framework. Finally, I will write about real-time adaptability and the need for production rules in this context.

8.1   Testing the Generalizability of my Findings

I strongly believe that this thesis represents an important step into the right direction. Nevertheless, due to the large amount of resources required to study the brain, neuroscientific research is usually limited to relatively small sample sizes. Additionally, the large variance in the data usually found in EEG studies (e.g. studies in **section 6.1** and **section 0**) hints at significant inter-individual differences. The situation is further complicated by findings that show that brain patterns can change over a short period of time. This becomes apparent, when one looks at the results from BCI studies that include more than one session. For instance, a study by Christensen, Estepp, Wilson, and Russell (2012) used a complex multitask to induce WML and found that classifier performance was significantly reduced by classifying across different days. Furthermore, previous research has shown that the type of modality used to induce mental states can have an impact on EEG signals used to infer the current user state. For example, a study by Mühl et al. (2011) used visual, auditory and audiovisual material to induce emotions and found opposing effects for visually and auditorily induced affective responses. The authors conclude that a classifier trained on one specific modality might poorly generalize to different modalities.

All this considerations make it difficult to assess in which settings one can rely on the findings from my as well as from other studies. This circumstance is especially

difficult in the case of the FAA, as my research already demonstrated that this feature is highly dependent on a specific context (compare study 1 in **section 6.1** and study 3 in **section 0**). Additionally, previous research has found that this laterality metric can be influenced by unilateral hand contractions (Harmon-Jones et al., 2010) as well as the individual seating position (Harmon-Jones, Gable, & Price, 2011). In this thesis I investigated two different emotion inductions (presentation of emotion inducing material in the first two studies and presentation of need-related affective situations in the third study) and found that only the presentation of a need-related situation produced an effect that could be found at the group level.

Therefore, I strongly recommend that my findings need to be further substantiated by validation studies that use different materials to induce complex mental states. Potential candidates are auditory material (e.g. using the IADS database; Bradley & Lang, 2007), audiovisual material (e.g. using the DEAP database; Koelstra & Mühl, 2012) or even more sophisticated approaches using virtual reality environments (e.g., Brouwer & Neerincx, 2011). The use of such different approaches for the induction of complex user states could shed additional light on the applicability and robustness of findings from previous neuroscience studies.

Additionally, I propose that future studies need to explicitly take the actual application setting into account. Such realistic scenarios are assumed to evoke more complex user states than normally addressed in controlled lab studies. Furthermore, previous work has shown that brain patterns can change due to the increased relevance in realistic scenarios (Mcdowell et al., 2013). The videogame industry seems to have already picked up on this idea, which is reflected in the development of EEG-based videogame controllers (Singer, 2008). In the case of a neurotutor, the most realistic setting would be the use of an actual ITS like the AutoTutor. To reach such an ambitious goal, I assume that one needs to develop and investigate methods that can be used in such realistic scenarios.

8.2 Other Analytical Approaches

As my analyses only focused on evoked frequency responses, I did not account for potential responses that were contained in the non-phase-locked part of the EEG signal. Approaches using induced activity like the ERD/S approach might reveal

additional aspects of the cognitive or affective responses in complex user states (Pfurtscheller & Lopes da Silva, 1999; Scharinger et al., 2015). However, to the best of my knowledge, nobody has yet compared evoked and induced approaches in the context of mental state detection. Furthermore, ERD/S approaches measure the activity relative to a preceding baseline. This might work well in a classical trial-based paradigm, but it is questionable how such approaches work in realistic environments where the user does not operate in repeated chunks of activity, but rather continuously. Even if such approaches might help to improve the approaches utilized in this dissertation, it appears implausible to assume that induced activity approaches do not suffer from context-sensitivity. Therefore, it is highly likely that one still has to search for other features that can be used to infer the user state during the use of a neurotutor.

8.3 The Need for new Features and Feature Extraction Methods

To be able to accurately detect a specific mental state via 'inverse inference', such EEG measures would need to be specific (Poldrack, 2006). Otherwise, one does not know if the state of interest is the cause of changes in a specific measure, or not, which makes the use of such measures for the detection of certain states highly problematic. However, most features used to infer the ongoing user state suffer from many-to-many relations (Cacioppo & Tassinary, 1990; Cacioppo, Tassinary, & Berntson, 2000). This means that multiple different processes can have an influence on a specific EEG measure. For instance, a study by Roy et al. (2013) induced fatigue and found that with increasing time-on-task the discriminability of WML was decreased. Especially in realistic scenarios that are assumed to evoke multiple different components of a complex user state such many-to-many relations could make it extremely difficult to detect specific user states. Furthermore, currently used classification approaches require training data for them to work. If the features used to infer a certain user state react different under each specific context than this would require that training data for each context is generated. The effort required for such an operation would probably outweigh potential gains promised via such systems by far. However, some lesser used approaches that do not only focus on band power modulations at specific brain locations might not suffer from this limitation. The next section will start with lesser utilized features that operate at the electrode level.

8.3.1 Feature Alternatives

Some researchers implemented approaches that are not as commonly used for the detection of mental states. For instance, a study by Aftanas, Reva, Varlamov, Pavlov, and Makhnev (2004) used the average energy over brain areas to assess emotions elicited by the presentation of emotional images. Their approach resulted in a complex feature space that consisted of six frequency bands for twelve (2 times 6 to account for hemispheric differences) different brain regions. The authors found a correlation between arousal and EEG responses in specific brain areas and frequency bands. Future studies would need to investigate if this approach is also sensible to changes in affective valence.

An even more exotic feature, called absolute logarithmic recoursing energy efficiency, was proposed by Murugappan, Ramachandran, and Yaacob (2010). This newly proposed feature outperformed other conventional features in a study that used an audiovisual emotion induction. Other, more widely used, approaches analyze the synchronization between different brain regions. Lorist et al. (2009), for example, induced fatigue via continuous task performance over two hours. After that they provided monetary reward as well as a positive social comparison to induce motivation. The authors found widespread coherence effects between regions of interest (following the approach by ten Caat, Maurits, & Roerdink, 2008) of mental fatigue which were independent from specific task manipulations. However, due to the effects of volume conduction an analysis of the band power could produce misleading results, therefore coherence analyses should only focus on phase information in EEG signals. Nonetheless, future research also needs to check these approaches for usability in affective valence detection.

Some analyses focus on modulations in the gamma band range to infer mental states (Makeig et al., 2011; Julie Onton & Makeig, 2009). For instance, a study by Müller, Keil, Gruber, & Elbert, 1999) used picture stimuli from the IAPS database and found right hemispheric gamma modulations in response to the affective valence stimulation. However, analyses of high frequency recordings can be quite challenging, due to the low amplitude of signals in this frequency range which seems to hardly pass the skull with sufficient energy. Additionally, muscle activity is known to produce

significant contributions to the scalp EEG in such frequency ranges (Whitham et al., 2007). This makes it difficult to separate cortical activity from artifactual activity due to muscle activity. EEG-based source localization approaches could help to identify high frequency responses in EEG data and separate them from artifacts.

8.3.2 Feature Extraction using advanced EEG-based Source Localization Methods

Unfortunately, the mapping of EEG signals to specific brain processes is usually ambiguous, meaning that EEG patterns evoked by different processes can look very similar. This is due to the fact that volume conduction spreads a specific brain signal across the whole cortex and therefore only superpositions of different processes can be recorded (Makeig et al., 1996). The third study of my thesis (see **section 0**) was already implementing source localization methods and thereby revealed effects that could not be found at the electrode level. Nonetheless, this should only be seen as initial attempt into a novel direction. Future studies need to acquire high-density EEG recordings with large data samples to obtain results that are more robust. Furthermore, the underdetermined nature of this approach requires that results be validated with other brain imaging methods. Ideally, this would include simultaneous recordings of EEG and fMRI as was done in a study by Wu, Eichele, and Calhoun (2010). However, to fully utilize the potential of such approaches, large datasets are necessary that allow large-scale analysis of EEG data. Efforts toward this goal are already underway, as reflected in projects like BigEEG (http://www.bigeeg.org/), the EEG Study Schema (http://www.eegstudy.org/) as well as the iNeuro project (https://mdcune.psych.ucla.edu/modules/ineuro; Bigdely-Shamlo, Makeig, & Robbins, 2016; Grisham, Lom, Lanyon, & L. Ramos, 2016). The analysis of such large datasets could also produce far more robust results and therefore help alleviate problems usually occurring in neuroscientific studies (see **section 8.1**).

While hopes are high for big data approaches in the context of EEG data, one should not forget about alternatives that already demonstrated some potential in the past. Most prominently the use peripheral signal used in combination with the EEG.

8.4   Peripheral Signals as a potential Addition to the EEG

The most commonly used biosignals in the context of user state detection are recordings of the heart rate, the electrodermal activity, respiration as well as the EEG (Broek et al., 2010). Heart rate sensors, like the photoplethysmograph, can be easy to deploy and are also cost efficient (Haag, Goronzy, Schaich, & Williams, 2004). A photoplethysmograph measures the blood flow (blood volume pulse) using a light source and a photo sensor which are attached to the skin. This signal can then be used to infer the heart-rate as well as the heart-rate-variability. Both have been shown to respond to changes in emotional as well as cognitive factors, however this unspecifity of the response makes it difficult to infer specific user states (Luay & Revett, 2011). Electrodermal activity (EDA) measures the conductivity of the skin via electrodes placed on the hands or feet, which is increased with higher perspiration (Broek et al., 2010; Koelstra et al., 2010). The EDA signal has been shown to be a good and sensitive indicator for stress, but, unfortunately, this signal is also influenced by other factors like outside temperature (Haag et al., 2004). Furthermore, the EDA could help reveal covert user states which even the user is not aware of, like stress responses (Norman, Mendolicchio, & Mordeniz, 2016). Respiration belts are tied around the chest and can record how deep and fast a person is breathing (Haag et al., 2004). Deep and slow breathing is usually related to relaxation, while shallow, quick and irregular respiration patterns are linked to anger or fear (Koelstra et al., 2010). However, breathing patterns also change with regard to other factors like olfactory stimulation (Homma & Masaoka, 2008).

Peripheral signals like the electrocardiogram, EDA and respiration were recorded in addition to the EEG in the emoback studies. Re-analysis of this data conducted as part of a bachelor thesis revealed that analyses of all three peripheral signal modalities achieved classification accuracies higher than seventy percent for the discrimination of affective valence (Bizu, 2016). However, due to potential time-confounds in the data, we did not further analyze these peripheral recordings.

Sensor fusion across different modalities appears to be an interesting approach which might be helpful in boosting classification performance. However, currently available evidence is ambiguous. While some studies find that sensor fusion using EEG and peripheral signals can improve classification performance (e.g., Chanel,

Kronegg, Grandjean, & Pun, 2006), other studies conclude that sensor fusion does not provide additional performance improvements (e.g., Hogervorst, Brouwer, & van Erp, 2014). Nonetheless, studies that include multiple different sensors could also have the benefit of allowing to compare different sensor approaches. This could enable the potential identification of circumstances where cost-effective sensor modalities (like heart-rate) provide similar results as high-end systems that are usually required for recordings of the EEG. Such potential reductions in hardware costs are an important issue if one wants to get technologies like neurotutors into classrooms. These ambitions are further supported by the development of cost-efficient EEG sensors, which will be described in the next section.

8.5 The Need for new EEG Recording Hardware

Making the technology required for the use of a neurotutor accessible is an essential goal if one wants to apply such approaches on a large scale. During my PhD I was already able to make first attempts in this direction. As part of two intramural research grant projects (LEAD IMRF: 19110516 and GriGerWa_523) we were able to investigate two EEG headsets and compare them with results obtained from clinical EEG systems. In the first project we were able to acquire a prototype of the Muse headband (http://www.choosemuse.com/) for validation purposes. Affective pictures were used to induce changes in affective valence in a pilot study. Our analyses found a peak accuracy of sixty-two percent, which was deemed insufficient for the use in a neurotutor. We concluded that the main factor for the poor performance was the low number (6) as well as the position of the electrodes (placed on the forehead as well as behind the ears, both areas are prone to artifact contamination). Accordingly, we decided to investigate a different headset in the second project, which would improve on these points. Therefore, we modified the Epoc headset (https://www.emotiv.com/epoc/) following the approach proposed by Debener, Minow, Emkes, Gandras, and de Vos (2012). The use of this EEG headset increased the amount of channels to fourteen and also allowed free placement of electrodes according to our requirements. Preliminary results have been very promising, achieving WML classification accuracies over seventy percent in real-time for a single

subject. However, future studies still need to further investigate the potential of this piece of hardware.

Nevertheless, previous research has already shown that cost-efficient EEG headsets can be used successfully in certain contexts (Pinegger, Wriessnegger, Faller, & Müller-Putz, 2016). After focusing primarily on technical considerations in the last sections, I will next focus on the proposed framework, which needs to be expanded and refined in future studies.

8.6   Expanding the proposed Framework

It is difficult to draw exact lines between different components of a complex user state. Therefore, one can argue about the assignment of different mental states to certain components of the proposed framework. An example is the case of attention, since all different components of the proposed framework require some form of attention. This is even more complicated since one can assume that most components exhibit strong interactions. For instance, affective processes can activate cognitive processes (e.g. worries due to negative emotions) and the other way around (e.g. negative emotions resulting from negative thoughts). However, even though the proposed framework is not able to account for all these complexities, it is a helpful framework to guide future research on the development of robust BCIs applications that are usable in settings outside the lab, as it explicitly assumes interactions between different components.

Nevertheless, future studies should also investigate other complex user states. Especially user states important in academic contexts like test anxiety, interest, motivation and flow. All those concepts appear to be highly relevant for a learning system like a neurotutor. In addition to extending the proposed framework and thereby adding additional complexity, one should consider application of previous results in a real-time scenario to see whether the findings hold under such conditions.

8.7   Real-time Application of Results in an Educational Context

The implementation of real-time detection of complex user states require different analytical approaches, as some offline preprocessing methods might be too slow to be used in a real-time scenario. Furthermore, some thought needs to be put on the

adjustments that are indicated when undesirable user states are present. This introduces the need for production rules that describe the exact behavior of such systems.

8.7.1 The Need for Production Rules to adapt towards the desired Learner State

Broadly speaking, adaptability of a given system in response to an undesirable state can be grouped into three categories (Gilleade, Dix, & Allanson, 2005). Firstly, systems can offer assistance if the user is stuck or unable to perform the task due to an inappropriate difficulty. Secondly, emotional elements can be incorporated into the display to promote positive emotions and avoid negative ones. For instance, encouragement and positive feedback as provided by the avatar implemented in the AutoTutor. Thirdly, the difficulty can be adjusted if the user experiences mental overload or to increase task engagement if the user is bored. The management of different levels of WML is the main objective of the cognitive load theory (CLT).

8.7.1.1 Cognitive Load Theory

According to CLT limitations in memory storage are the main determinant for learning in complex, real-world scenarios, and therefore act as a bottleneck for information processing (Sweller, Van Merrienboer, & Paas, 1998). Hence, CLT puts a special emphasis on the careful management of limited WM capacity (Cowan, 2014). In CLT WML is divided into three categories: intrinsic, extraneous and germane load.
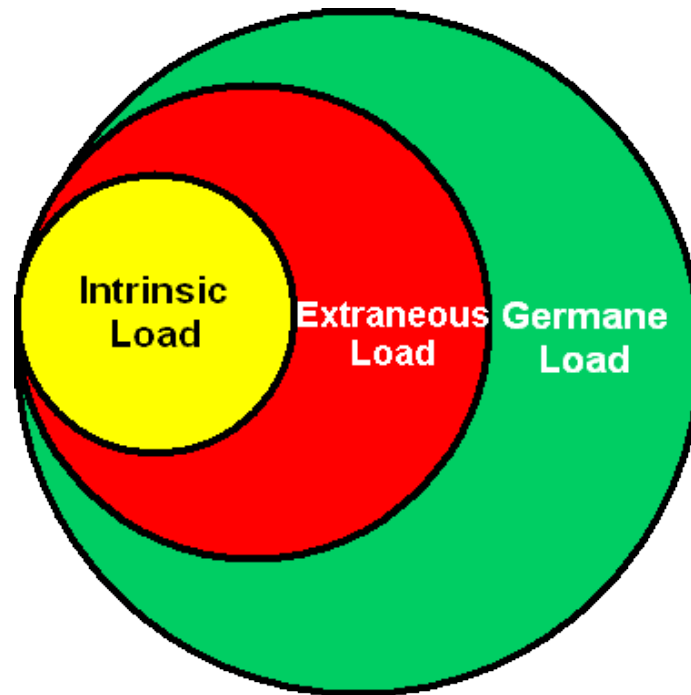
Figure 49. Different types of cognitive load according to cognitive load theory (http://www.simulations.co.uk/images/Cognitive%20Load%20Types.gif)**.**

According to CLT, intrinsic load is directly related to the specific task at hand. Extraneous load is related to the way in which a specific task is presented and can be optimized via instructional design. Germane load is used for the processing, construction and automation of schemas and therefore beneficial for the learning progress. CLT is widely used in educational research, especially in the context of instructional design (Sweller & Sweller, 1994), where the main goal is to reduce extraneous load and promote germane load. WM as seen in CLT is conceptualized based on the MCM by Baddeley (2000; see **section 3.1.1.3**). Previously, CLT has already been used with the aim to induce a state which could be interesting in the context of neurotutors, called flow (Chanel, 2009).

8.7.1.2 Adapting a Learning Environment to induce a Flow State

Although education is a vital part of modern society, not all learning experiences are joyful. A study by Csikszentmihalyi and Larson (1987) found that, on average, students feel bored during thirty-two percent of classroom instruction time. This can lead to impaired performance as well as reduced motivation. Adaptive learning environments like ITS could help to alleviate this problem by creating individualized learning environments that keep the learner in the zone of proximal development and thereby

improve the learning process (Kirschner & Gerjets, 2006). One very simple framework which could allow for such an adaption would be the flow channel proposed by (Csikszentmihalyi, 1990). Flow is a state in which one is engaged in a task that challenges just with the optimal intensity (Csiksczentmihalyi et al., 2004). People call it 'being in the zone', 'rapture' or 'ecstasy'. It is distinct from mere pleasure by an active component. While one can find pleasure in passive activities like watching TV or doing drugs, the flow state requires conscious will and an eagerness to grow. It is goal directed and gives a sense of mastery, which makes the flow state an optimal target state for adaptive learning. The flow channel tries to induce this state by simply changing the difficulty of the task. If the task is too difficult for the learner the system presents tasks with a lower difficulty and thereby tries to avoid feelings of anxiety. If the tasks are not difficult enough (e.g., due to changes in the learners´ competence) the system switches to more challenging tasks to avoid feelings of boredom (see **Figure 50**).
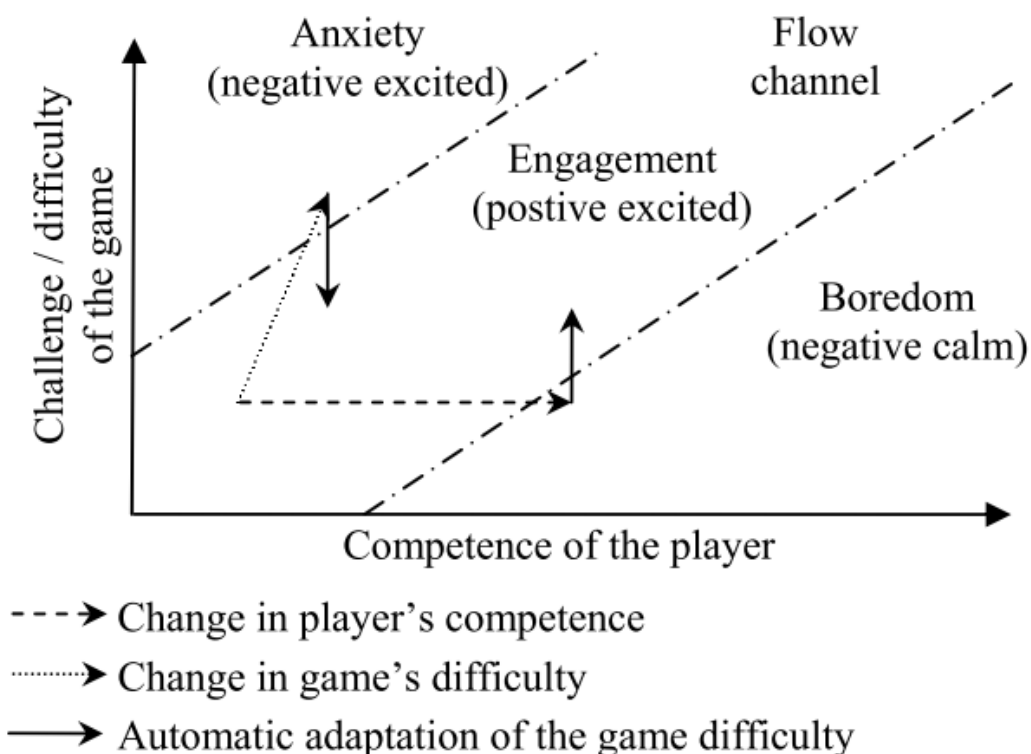
Figure 50. The flow channel (Csikszentmihalyi, 1990).

While this simple framework provides a viable starting point, one could argue that changes in task difficulty alone will not suffice to induce a complex user state like flow, as it presumably consists of multiple different basic mental states. The flow channel only tracks the affective components of the ongoing user state. Cognitive components are influenced by changes of task difficulty, but not actively tracked through this framework. The three-dimensional framework I proposed in this dissertation would allow to track cognitive as well as affective aspects that are assumed to be part of the flow state. Furthermore, the framework would also allow to account for additional context factors that could potentially influence system performance. While the flow channel only offers limited ways to induce a desired target state, more sophisticated approaches have already been developed for existing ITS (D'Mello et al., 2008). Such approaches, however, would need to be adapted for the use in neurotutors. While the previous section focused on all the work that still needs to be done before we can all profit from the research in this field, the next section will try to take a look into the distant future and illustrate the enormous potential of such technologies in the educational context.

9 The potential Use of Findings in the Educational Context

Due to the innovative and inherently interdisciplinary nature of this research project there was a lot of risk involved. To the best of my knowledge, to date no framework that tries to describe complex user states in the context of EEG-based human-machine interaction was developed. Even though a lot of further research is still required, it seems likely that the research in this field could profit from a broader view on user states that includes different aspects.

In the future, I see at least three different contexts for potential applications of the findings from this thesis. The first context relates to the research conducted in the field of brain-based human-machine interaction. I'm convinced that it is necessary to test experimental approaches and methods under more realistic conditions if one wants to produce findings that are robust in settings outside the lab. This means that we have to identify the most relevant user states for a specific target scenario and then optimize paradigms and classification approaches for this specific context.

The second context relates to the application of brain-based learner state detection in the context of macro adaption to optimize learning material for specific student types. Once, we are able to develop systems that can reliably detect certain learner states, then we can potentially identify which learning material is suited for what type of students. One can imagine a case where additional learning material in the form of animations could benefit or hurt learners. If a certain learner has sufficient cognitive capacities to process such animations without drawing mental resources from the primary task than this would probably help to facilitate memory encoding. If a learner does not have such additional resources available then the presentation of additional material could be detrimental to the learning progress. Since introspection is a highly unreliable method, the use of brain-based detection methods seems very promising in such a context. Additionally, the reliable detection of different user states during learning could identify factors that are associated with high and low achievement on an individualized level that is not possible with currently applied methods like large-scale questionnaires.

The third and final context relates to the real-time adaptation in the form of a Neurotutor. Neurotutors could provide individualized learning environments that challenge those gifted while supporting those lacking capabilities. Thus, once

expenses are reduced through the introduction of cost-efficient hardware, we could be able to provide ever more members of society with high-quality education currently only accessible for those with a wealthy background. Furthermore, individual tracking of the learning progress could be used to identify deviations from an optimal learning course early, thereby potentially avoiding significant drops in key factors like motivation and thereby potentially preventing detrimental cognitive as well as affective developments. Newly adopted learning schemes like flip teaching (Mason, Shuman, & Cook, 2013) and MOOC (Yuan & Powell, 2013) are strongly dependent on self-regulation. Neurotutors have already been shown to be able to support users in this context by providing adaptability that can optimize the learning process (see **section 3.1**). In the future one can imagine the widespread availability of free education, which currently too many people on this planet are lacking.

If such systems are implemented at a large scale, the use of these systems could create large datasets that could be analyzed using big data approaches. This could potentially reveal additional aspects of the learning process that are not accessible via other means, thereby further improving the quality of the teaching material. However, it has to be noted that the collection of this kind of personal data also forces us to think about the ethical implications and the potential misuse of such data. Issues that will become more and more important as we venture deeper into the digital age.

As a closing remark I want to state that research that is channeled in this direction holds the promise of creating immensely powerful tools. However, with great power… comes great responsibility!

10 References

Aftanas, L. I., Reva, N. V, Varlamov, A. A., Pavlov, S. V, & Makhnev, V. P. (2004). Analysis of evoked EEG synchronization and desynchronization in conditions of emotional activation in humans: temporal and topographic characteristics. *Neuroscience and Behavioral Physiology, 34(8)*, 859–867. http://doi.org/10.1023/B:NEAB.0000038139.39812.eb

Ahern, G. L., & Schwartz, G. E. (1985). Differential lateralization for positive and negative emotion in the human brain: EEG spectral analysis. *Neuropsychologia, 23(6)*, 745–755. http://doi.org/10.1016/0028-3932(85)90081-8

Albanese, P. J. (1987). The nature of preferences: An exploration of the relationship between economics and psychology. *Journal of Economic Psychology, 8(1)*, 3–18. http://doi.org/10.1016/0167-4870(87)90003-1

Alguacil, S., Tudela, P., & Ruz, M. (2013). Cognitive and affective control in a flanker word task: Common and dissociable brain mechanisms. *Neuropsychologia, 51(9)*, 1663–1672. http://doi.org/10.1016/j.neuropsychologia.2013.05.020

Allen, J. J. B., Coan, J. a, & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. *Biological Psychology, 67(1–2)*, 183–218. http://doi.org/10.1016/j.biopsycho.2004.03.007

Allison, B. Z., Brunner, C., Altstätter, C., Wagner, I. C., Grissmann, S., & Neuper, C. (2011). A hybrid ERD / SSVEP BCI for continuous simultaneous two dimensional cursor control. *Journal of Neuroscience Methods, 209(2)*, 299-307. http://dx.doi.org/10.1016/j.jneumeth.2012.06.022

Anderson, J., & Corbett, A. (1995). Cognitive tutors: Lessons learned. *The journal of the learning sciences*, *4*(2), 167-207. http://dx.doi.org/10.1207/s15327809jls0402_2

Anderson, K. (1986). Computer-assisted instruction. *Journal of Medical Systems, 10(2)*, 163–171. http://doi.org/10.1007/BF00993122

Ansari, D., & Coch, D. (2006). Bridges over troubled waters: education and cognitive neuroscience. *Trends in Cognitive Sciences, 10(4)*, 146–51. http://doi.org/10.1016/j.tics.2006.02.007

Ansari, D., Smedt, B., & Grabner, R. H. (2011). Neuroeducation – A critical overview of an emerging field. *Neuroethics, 5(2)*, 105–117. http://doi.org/10.1007/s12152-011-9119-3

Ayaz, H., Izzetoglu, M., Bunce, S., Heiman-Patterson, T., & Onaral, B. (2007). Detecting cognitive activity related hemodynamic signal for brain computer interface using functional near infrared spectroscopy. *3rd International*

IEEE/EMBS Conference on Neural Engineering, 342–345. http://doi.org/10.1109/CNE.2007.369680

Azcarraga, J., Marcos, N., & Suarez, M. T. (2014). Modelling EEG Signals for the prediction of academic emotions. In *Workshop on Utilizing EEG Input in Intelligent Tutoring Systems (ITS2014 WSEEG)*

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4(11)*, 417–423. http://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4(10)*, 829–839. http://doi.org/10.1038/nrn1201

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63(1)*, 1–29. http://doi.org/10.1146/annurev-psych-120710-100422

Baldwin, C. L., & Penaranda, B. N. (2012). Adaptive training using an artificial neural network and EEG metrics for within- and cross-task workload classification. *NeuroImage, 59(1)*, 48–56. http://doi.org/10.1016/j.neuroimage.2011.07.047

Belitski, A., Farquhar, J., & Desain, P. (2011). P300 audio-visual speller. *Journal of Neural Engineering, 8*, 25022. http://doi.org/10.1088/1741-2560/8/2/025022

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7(6)*, 1129–1159. http://doi.org/10.1162/neco.1995.7.6.1129

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience,* 87(1): 527-570. http://doi.org/10.1007/BF01797193

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T.,… & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine, 78(5 Suppl)*, B231-44.

Bigdely-Shamlo, N., Makeig, S., & Robbins, K. A. (2016). Preparing laboratory and real-world EEG data for large-scale analysis: A containerized approach. *Frontiers in Neuroinformatics, 10(March)*, 7. http://doi.org/10.3389/fninf.2016.00007

Birbaumer, N. (2006). Breaking the silence: Brain-computer interfaces (BCI) for communication and motor control. *Psychophysiology, 43(6)*, 517–532. http://doi.org/10.1111/j.1469-8986.2006.00456.x

Bizu, S. N. (2016). Klassifikation emotionaler Zustände. Bachelor thesis. Eberhard Karls Universität Tübingen.

Blankertz, B., Lemm, S., Treder, M., Haufe, S., & Müller, K.-R. (2011). Single-trial analysis and classification of ERP components--a tutorial. *NeuroImage, 56(2)*, 814–825. http://doi.org/10.1016/j.neuroimage.2010.06.048

Bowers, J. S. (2016). Psychology, not educational neuroscience, is the way forward for improving educational outcomes for all children: Reply to Gabrieli (2016) and Howard-Jones et al. (2016). *Psychological Review, 123(5)*, 628–635. http://doi.org/10.1037/rev0000043

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

Bradley, M. M., & Lang, P. J. (2007). The International Affective Digitized Sounds (IADS-2): Affective ratings of sounds and instruction manual. *University of Florida, Gainesville, FL, Tech. Rep. B-3*.

Brandl, S., Frølich, L., Höhne, J., Müller, K.-R., & Samek, W. (2016). Brain-computer interfacing under distraction: an evaluation study. *Journal of Neural Engineering, 13(5)*, 56012. http://doi.org/10.1088/1741-2560/13/5/056012

Briesemeister, B., Tamm, S., Heine, A., & Jacobs, A. (2013). Approach the good, withdraw from the bad - A review on frontal alpha asymmetry measures in applied psychological research. *Psychology, 4(3),* 261–267. http://doi.org/10.4236/psych.2013.43A039

Briggs, K. E., & Martin, F. H. (2009). Affective picture processing and motivational relevance: Arousal and valence effects on ERPs in an oddball task. *International Journal of Psychophysiology, 72(3)*, 299–306. http://doi.org/10.1016/j.ijpsycho.2009.01.009

Broek, E. van den, Lisý, V., Janssen, J. H., Westerink, J. H. D. M., Schut, M., & Tuinenbreijer, K. (2010). Affective man-machine interface: Unveiling human emotions through biosignals. In *Biomedical Engineering Systems and Technologies* (pp. 21–47).

Brouwer, A. M., Neerincx, M. A., Kallen, V. L., van der Leer, L., & ten Brinke, M. (2011). EEG alpha asymmetry, heart rate variability and cortisol in response to virtual reality induced stress. *Journal of Cybertherapy & Rehabilitation*, *4*(1), 21-34.

Brouwer, A.-M., Hogervorst, M., van Erp, J. B. F., Heffelaar, T., Zimmerman, P. H., & Oostenveld, R. (2012). Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of Neural Engineering, 9, 45008*. http://doi.org/10.1088/1741-2560/9/4/045008

Brunner, C., Naeem, M., Leeb, R., Graimann, B., & Pfurtscheller, G. (2007). Spatial filtering and selection of optimized components in four class motor imagery

EEG data using independent components analysis. *Pattern Recognition Letters*, *28*(8), 957-964. http://doi.org/10.1016/j.patrec.2007.01.002

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4(6)*, 215–222. http://dx.doi.org/10.1016/S1364-6613(00)01483-2

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45(1)*, 16–28. http://doi.org/10.1037/0003-066X.45.1.16

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). Psychophysiological science. In *Handbook of Psychophysiology* (pp. 3–22).

Carbonell, J. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man Machine Systems, 11(4)*, 190–202. http://doi.org/10.1109/TMMS.1970.299942

Carretié, L. (2014). Exogenous (automatic) attention to emotional stimuli: a review. *Cognitive, affective & behavioral neuroscience (Vol. 14).* http://doi.org/10.3758/s13415-014-0270-2

Chambers, J. A., & Sprecher, J. W. (1983). Computer-assisted instruction: its use in the classroom. Prentice-Hall Direct.

Chanel, G. (2009). Emotion assessment for affective computing based on brain and peripheral signals. Doctoral Thesis. University of Geneva.

Chanel, G., Kronegg, J., Grandjean, D., & Pun, T. (2006). Emotion assessment : Arousal evaluation using EEG's and peripheral physiological signals. *Lecture Notes in Computer Science, 4105*, 530–537. http://doi.org/10.1007/11848035_70

Christensen, J. C., Estepp, J. R., Wilson, G. F., & Russell, C. A. (2012). The effects of day-to-day variability of physiological data on operator functional state classification. *NeuroImage, 59(1)*, 57–63. http://doi.org/10.1016/j.neuroimage.2011.07.091

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3(3)*, 215–229. http://doi.org/10.1038/nrn755

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197-223. http://doi.org/10.1007/s10648-013-9246-y

Csiksczentmihalyi, M., Kolo, C., & Baur, T. (2004). Flow: The psychology of optimal experience. *Australian Occupational Therapy Journal*, 1–6.

Csikszentmihalyi, M. (1990). Flow: The psychology of optimal experience. NY: Cambridge University Press. http://doi.org/10.5465/AMR.1991.4279513

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of Nervous and Mental Disease*, 175(9), 526–36.

D'Mello, S., Jackson, T., Craig, S., Morgan, B., Chipman, P., White, H.,… & Graesser, A. (2008). AutoTutor detects and responds to learners affective and cognitive states. *Proceedings of the Workshop on Emotional and Cognitive Issues in ITS in Conjunction with the 9th International Conference on Intelligent Tutoring Systems.*

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153–170. http://doi.org/10.1016/j.learninstruc.2012.05.003

D'Mello, S., Picard, R. W., & Graesser, A. (2007). Toward an Affect-Sensitive AutoTutor. *IEEE Intelligent Systems, 22(4)*, 53–61. http://doi.org/10.1109/MIS.2007.79

Dalgleish, T. (2004). The emotional brain. *Nature Reviews Neuroscience, 5(7)*, 583–9. http://doi.org/10.1038/nrn1432

Daly, J. J., & Wolpaw, J. R. (2008). Brain-computer interfaces in neurological rehabilitation. *The Lancet Neurology, 7(11)*, 1032–1043. http://doi.org/10.1016/S1474-4422(08)70223-0

Davidson, R. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition, 151*, 125–151. http://dx.doi.org/10.1016/0278-2626(92)90065-T

Davidson, R. J., & Tomarken, A. J. (1989). Laterality and emotion: An electrophysiological approach. *Handbook of neuropsychology*, 3, 419-441.

Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Doss, R. C. (1992). Individual differences in anterior brain asymmetry and fundamental dimensions of emotion. *Journal of personality and social psychology*, *62*(4), 676. http://dx.doi.org/10.1037/0022-3514.62.4.676

Debener, S., Minow, F., Emkes, R., Gandras, K., & de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology, 49*, 1617–1621. http://doi.org/10.1111/j.1469-8986.2012.01471.x

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods. 134(1)*, 9-21. http://dx.doi.org/10.1016/j.jneumeth.2003.10.009

Delorme, A., Mullen, T., Kothe, C., Acar, Z. A., Bigdely-Shamlo, N., Vankov, A., & Makeig, S. (2011). EEGLAB, SIFT, NFT, BCILAB, and ERICA: new tools for

advanced EEG processing. *Computational Intelligence and Neuroscience, 2011, 130714.* http://dx.doi.org/10.1155/2011/130714

Delorme, A., Palmer, J., Onton, J., Oostenveld, R., & Makeig, S. (2012). Independent EEG sources are dipolar. *PloS One, 7(2)*, e30135. http://doi.org/10.1371/journal.pone.0030135

Demaree, H. A., Everhart, D. E., Youngstrom, E. A., & Harrison, D. W. (2005). Brain lateralization of emotional processing: historical roots and a future incorporating "dominance". *Behavioral and Cognitive Neuroscience Reviews, 4*, 3–20. http://doi.org/10.1177/1534582305276837

Diamond, A. (2013). Executive functions. *The Annual Review of Psychology, 64*, 135–168. http://doi.org/10.1146/annurev-psych-113011-143750

Ding, C., & He, X. (2004). K -means clustering via principal component analysis. In *Twenty-first international conference on Machine learning - ICML '04 (p. 29)*. New York, New York, USA: ACM Press. http://doi.org/10.1145/1015330.1015408

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International journal of science education*, *25*(6), 671-688. http://doi.org/10.1080/09500690305016

Duvinage, M., Castermans, T., Petieau, M., Hoellinger, T., Cheron, G., & Dutoit, T. (2013). Performance of the Emotiv Epoc headset for P300-based applications. *Biomedical Engineering Online, 12(1)*, 56. http://doi.org/10.1186/1475-925X-12-56

Ekman, P., Friesen, W. V, O'Sullivan, M., Diacoyanni-Tarlatzis, I., Krause, R., Pitcairn, T., … & Tzavaras, A. (1987). Universals and cultural differences in facial expressions of emotion. *Journal of Personality and Social Psychology, 53(4)*, 712–717. http://dx.doi.org/10.1037/0022-3514.53.4.712

Ellis, H., & Ashbrook, P. (1988). Resource allocation model of the effects of depressed mood states on memory. In *Affect, cognition, and social behavior* (pp. 25–43). Lewiston, NY: Hogrefe.

Ellis, H. C., & Ashbrook, P. W. (1989). The "State" of Mood and Memory Research: A selective review. *Journal of Social Behavior and Personality, 4(2),* 1-21.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11(1)*, 19–23. http://doi.org/10.1111/1467-8721.00160

Fairclough, S. H. (2009). Fundamentals of physiological computing. *Interacting with Computers, 21(1–2)*, 133–145. http://doi.org/10.1016/j.intcom.2008.10.011

Fairclough, S. H., & Roberts, J. S. (2011). Effects of performance feedback on cardiovascular reactivity and frontal EEG asymmetry. *International Journal of Psychophysiology, 81(3)*, 291–298. http://doi.org/10.1016/j.ijpsycho.2011.07.012

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51(2-3)*, 87–107. http://doi.org/10.1016/S0301-0511(99)00031-9

Faller, J., Vidaurre, C., Solis Escalante, T., Neuper, C., & Scherer, R. (2012). Autocalibration and recurrent adaptation: Towards a plug and play online ERD-BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society, XX(X), 1–8.* http://doi.org/10.1109/TNSRE.2012.2189584

Fraser, K., Ma, I., Teteris, E., Baxter, H., Wright, B., & Mclaughlin, K. (2012). Emotion, cognitive load and learning outcomes during simulation training. *Medical Education, 46(11)*, 1055-1062. http://doi.org/10.1111/j.1365-2923.2012.04355.x

Frey, J., Mühl, C., Lotte, F., & Hachet, M. (2014). Review of the use of electroencephalography as an evaluation method for human-computer interaction. *PhyCS 2014 - Proceedings of the International Conference on Physiological Computing Systems*, 214–223. http://doi.org/10.5220/0004708102140223

Frijda, N. H. (1994). Varieties of affect: Emotions and episodes, moods, and sentiments. In *The Nature of Emotion: Fundamental Questions* (pp. 197–202). Oxford University Press.

Galway, L., McCullagh, P., Lightbody, G., Brennan, C., & Trainor, D. (2015). The potential of the Brain-Computer Interface for learning: A technology review. *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 1554–1559. http://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.234

Galy, E., Cariou, M., & Mélan, C. (2012). What is the relationship between mental workload factors and cognitive load types? *International Journal of Psychophysiology, 83(3)*, 269-275. http://doi.org/10.1016/j.ijpsycho.2011.09.023

Gerjets, P., Walter, C., Rosenstiel, W., Bogdan, M., & Zander, T. O. (2014). Cognitive state monitoring and the design of adaptive instruction in digital environments: lessons learned from cognitive workload assessment using a passive brain-computer interface approach. *Frontiers in Neuroscience, 8(December)*, 1–21. http://doi.org/10.3389/fnins.2014.00385

Gerrards-Hesse, A., Spies, K., & Hesse, F. W. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology, 85*, 55–78. http://doi.org/10.1111/j.2044-8295.1994.tb02508.x

Gevins, A., Leong, H., Du, R., Smith, M. E., Le, J., DuRousseau, D.,… & Libove, J. (1995). Towards measurement of brain function in operational environments. *Biological Psychology, 40(1–2)*, 169–186. http://dx.doi.org/10.1016/0301-0511(95)05105-8

Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex (New York, N.Y.: 1991), 10(9)*, 829–839. http://doi.org/10.1093/cercor/10.9.829

Gilleade, K., Dix, A., & Allanson, J. (2005). Affective videogames and modes of affective gaming: assist me, challenge me, emote me, 1–7. In *The 2005 International Conference on Changing Views: Worlds in Play*

Graesser, A. (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research, 1(1)*, 35–51. http://doi.org/10.1016/S1389-0417(99)00005-4

Graesser, A., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: a tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc, 36(2)*, 180–192. http://doi.org/10.3758/BF03195563

Grisham, W., Lom, B., Lanyon, L., & L. Ramos, R. (2016). Proposed training to meet challenges of large-scale data in neuroscience. *Frontiers in Neuroinformatics, 10(July)*, 1–6. http://doi.org/10.3389/fninf.2016.00028

Grissmann, S., Faller, J., Spüler, M., Zander, T. O., Kelava, A., Scharinger, C., & Gerjets, P. (under review). Electroencephalography based identification of working memory load and affective valence in an N-back task with emotional stimuli. *PLoS ONE.*

Grissmann, S., Solis-Escalante, T., Brunner, C., Scherer, R., Neuper, C., & Müller-Putz, G. R. (2012). How to train a hybrid BCI based on ERD and SSVEP final. Poster at the 2012 BBCI Workshop. Berlin

Grissmann, S., Zander, T. O., Faller, J., Brönstrup, J., Kelava, A., Gramann, K., & Gerjets, P. (under review). Brain-computer interfaces for natural environments require context-aware user state models: An illustration based on a loss of control paradigm. *Frontiers in Human Neuroscience.*

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

Gruber, T., Tsivilis, D., Giabbiconi, C.-M., & Müller, M. M. (2008). Induced electroencephalogram oscillations during source memory: familiarity is reflected

in the gamma band, recollection in the theta band. *Journal of Cognitive Neuroscience, 20*, 1043–1053. http://doi.org/10.1162/jocn.2008.20068

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2011). Leaning embodies desire: Evidence that leaning forward increases relative left frontal cortical activation to appetitive stimuli. *Biological Psychology, 87(2)*, 311–313. http://doi.org/10.1016/j.biopsycho.2011.03.009

Harmon-Jones, E., Gable, P., & Peterson, C. K. (2010). The role of asymmetric frontal cortical activity in emotion-related phenomena: a review and update. *Biological Psychology, 84(3)*, 451–62. http://doi.org/10.1016/j.biopsycho.2009.08.010

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.   http://dx.doi.org/10.1016/S0166-4115(08)62386-9

Hembree, R. (1988). Correlates, causes, effects , and treatment of test anxiety. *Review of educational research, 58(1)*, 47–77. http://dx.doi.org/10.2307/1170348

Hendelman, W. (2005). Atlas of functional neuroanatomy (2nd ed.). Taylor & Francis Group.

Heraz, A., & Frasson, C. (2007). Predicting the three major dimensions of the learner's emotions from brainwaves. *International Journal of Computer Science, 25(November)*, 323–329.

Hettich, D. T., Bolinger, E., Matuz, T., Birbaumer, N., & Spüler, M. (2016). EEG Responses to Auditory Stimuli for Automatic Affect Recognition. *Frontiers in Neuroscience, 10(244)*. http://doi.org/10.3389/fnins.2016.00244

Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion (Washington, D.C.), 11(2)*, 346–353. http://doi.org/10.1037/a0022771

Hogervorst, M. A., Brouwer, A. M., & van Erp, J. B. F. (2014). Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload. *Frontiers in Neuroscience, 8(Oct)*, 1–14. http://doi.org/10.3389/fnins.2014.00322

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109(4)*, 679–709. http://doi.org/10.1037//0033-295X.109.4.679

Huang, D., Zhang, H., Ang, K., Guan, C., Pan, Y., Wang, C., & Yu, J. (2012, March). Fast emotion detection from EEG using asymmetric spatial filtering. In *2012*

IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 589-592). http://doi.org/10.1109/ICASSP.2012.6287952

Huang, Y.-M., Baddeley, A., & Young, A. W. (2008). Attentional capture by emotional stimuli is modulated by semantic processing. *Journal of Experimental Psychology: Human Perception and Performance, 34(2)*, 328–39. http://doi.org/10.1037/0096-1523.34.2.328

Jacobs, J., Hwang, G., Curran, T., & Kahana, M. J. (2006). EEG oscillations and recognition memory: Theta correlates of memory retrieval and decision making. *NeuroImage, 32(2)*, 978–987. http://doi.org/10.1016/j.neuroimage.2006.02.018

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N -back task as a working memory measure. *Memory, 18(4)*, 394–412. http://doi.org/10.1080/09658211003702171

Jatzev, S., Zander, T., Filippis, M. De, & Kothe, C. (2008). Examining causes for non-stationarities: The loss of controllability is a factor which induces non-stationarities. In *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course 2008* (pp. 138–143). Graz, Austria.

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Neuroscience, 15(8)*, 1395–1399. http://doi.org/10.1046/j.1460-9568.2002.01975.x

Jurado, M., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*(3), 213-233. http://doi.org/10.1007/s11065-007-9040-z

Kaminski, J., Brzezicka, A., & Wrobel, A. (2011). Short-term memory capacity (7 ± 2) predicted by theta to gamma cycle length ratio. *Neurobiology of Learning and Memory, 95(1)*, 19-23. http://doi.org/10.1016/j.nlm.2010.10.001

Kappenman, E. S., & Luck, S. J. (2012). The Oxford Handbook of Event-Related Potential Components. http://doi.org/10.1093/oxfordhb/9780195374148.001.0001

Kim, M.-K., Kim, M., Oh, E., & Kim, S.-P. (2013). A review on the computational methods for emotional state estimation from the human EEG. *Computational and Mathematical Methods in Medicine, 2013, 573734*. http://doi.org/10.1155/2013/573734

Kim, Lee, M., Chung, Y., & Bong, M. (2010). Comparison of brain activation during norm-referenced versus criterion-referenced feedback: The role of perceived competence and performance-approach goals. *Contemporary Educational Psychology, 35(2)*, 141–152. http://doi.org/10.1016/j.cedpsych.2010.04.002

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*, 352–358. http://doi.org/10.1037/h0043688

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. http://doi.org/119004

Kirschner, P., & Gerjets, P. (2006). Instructional design for effective and enjoyable computer-supported learning. *Computers in Human Behavior, 22(1)*, 1–8. http://doi.org/10.1016/j.chb.2005.01.004

Kleih, S., Nijboer, F., Halder, S., & Kübler, A. (2010). Motivation modulates the P300 amplitude during brain-computer interface use. *Clinical Neurophysiology, 121(7)*, 1023–1031. http://doi.org/10.1016/j.clinph.2010.01.034

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology, 24(1–2)*, 61–100. http://dx.doi.org/10.1016/S0167-8760(96)00057-8

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2), 169-195. http://dx.doi.org/10.1016/S0165-0173(98)00056-3

Klimesch, W., Freunberger, R., Sauseng, P., & Gruber, W. (2008). A short review of slow phase synchronization and memory: Evidence for control processes in different memory systems? *Brain Research, 1235*, 31–44. http://doi.org/10.1016/j.brainres.2008.06.049

Klingberg, T. (2009). The overflowing brain information overload and the limits of working memory. Oxford University Press.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19(3)*, 239–264. http://doi.org/10.1007/s10648-007-9049-0

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30-43.

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T.,... & Patras, I. (2012). Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, *3*(1), 18-31. http://doi.org/10.1109/T-AFFC.2011.15

Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K.-R., Curio, G., … Kincses, W. E. (2007). Improving human performance in a real operating environment through real-time mental workload detection. In *G. Dornhege, J.*

del R. Millán, T. Hinterberger, D. J. McFarland, & K.-R. Müller (Eds.), Toward Brain-Computer Interfacing (pp. 409–422). MIT Press.

Kop, W. J., Synowski, S. J., Newell, M. E., Schmidt, L. A., Waldstein, S. R., & Fox, N. A. (2011). Autonomic nervous system reactivity to positive and negative mood induction: The role of acute psychological responses and frontal electrocortical activity. *Biological Psychology, 86(3)*, 230–238. http://doi.org/10.1016/j.biopsycho.2010.12.003

Kopf, J., Dresler, T., Reicherts, P., Herrmann, M. J., & Reif, A. (2013). The effect of emotional content on brain activation and the late positive potential in a word n-back task. *PLoS ONE, 8(9)*, 1–11. http://doi.org/10.1371/journal.pone.0075598

Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X., & Milhama, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. *NeuroImage, 32(1)*, 477–484. http://doi.org/10.1016/j.neuroimage.2006.02.047

Krauledat, M., Tangermann, M., Blankertz, B., & Müller, K.-R. (2006). Towards Zero Training for Brain-Computer Interfacing. *PLoS ONE, 3(8)*. http://dx.doi.org/10.1371/journal.pone.0002967

Krause, C. M. (2003). Brain electric oscillations and cognitive processes. In *Experimental Methods in Neuropsychology* (pp. 111–130). Springer US. http://doi.org/10.1007/978-1-4615-1163-2_6

Krause, C. M., Sillanmäki, L., Koivisto, M., Saarela, C., Häggqvist, A., Laine, M., & Hämäläinen, H. (2000). The effects of memory load on event-related EEG desynchronization and synchronization. *Clinical Neurophysiology, 111(11)*, 2071–2078. http://doi.org/10.1016/S1388-2457(00)00429-6

Krumpe, T., Walter, C., Rosenstiel, W., & Spüler, M. (2016). Asynchronous P300 classification in a reactive brain-computer interface during an outlier detection task. *Journal of Neural Engineering, 13(4)*, 46015. http://doi.org/10.1088/1741-2560/13/4/046015

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A comparison of classification techniques for the P300 Speller. *Journal of Neural Engineering, 3(4)*, 299–305. http://doi.org/10.1088/1741-2560/3/4/007

Kübler, A., Neumann, N., Kaiser, J., Kotchoubey, B., Hinterberger, T., & Birbaumer, N. (2001). Brain-computer communication: self-regulation of slow cortical potentials for verbal communication. *Archives of Physical Medicine and Rehabilitation, 82(11)*, 1533–1539. http://dx.doi.org/10.1053/apmr.2001.26621

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L.,… & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10(3)*, 120–131. http://dx.doi.org/10.1002/1097-0193(200007)10:3<120::AID-HBM30>3.0.CO;2-8

Lang, P. J., Bradley, M., & Cuthbert, B. (1997). International Affective Picture System (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention*, 39–58. http://doi.org/10.1027/0269-8803/a000147

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *Gainsville, USA: University of Florida, International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual*. Technical Report A-6.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30(3)*, 261–73. http://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Larkin, J. H. E., & Chabay, R. W. E. (1992). Computer-Assisted Instruction and Intelligent Tutoring Systems: Shared goals and complementary approaches. *Technology in Education Series*. Lawrence Erlbaum Associates, Inc., Publishers, 365 Broadway, Hillsdale, NJ 07462 (hardback: ISBN-0-8058-0232-0).

Law, J., & Jolliffe, I. T. (1987). Principal Component Analysis. *The Statistician, 36(4)*, 432. http://doi.org/10.2307/2348864

Leech, R., Kamourieh, S., Beckmann, C. F., & Sharp, D. J. (2011). Fractionating the default mode network: distinct contributions of the ventral and dorsal posterior cingulate cortex to cognitive control. *The Journal of Neuroscience, 31(9)*, 3217–3224. http://doi.org/10.1523/JNEUROSCI.5626-10.2011

Levenson, R. W. (1999). The intrapersonal functions of emotion. *Cognition & Emotion, 13(5)*, 481–504. http://doi.org/10.1080/026999399379159

Leventhal, A. M., Martin, R. L., Seals, R. W., Tapia, E., & Rehm, L. P. (2007). Investigating the dynamics of affect: Psychological mechanisms of affective habituation to pleasurable stimuli. *Motivation and Emotion, 31(2)*, 145–157. http://doi.org/10.1007/s11031-007-9059-8

Li, Y., Li, X., Ratcliffe, M., Liu, L., Qi, Y., & Liu, Q. (2011). A real-time EEG-based BCI system for attention recognition in ubiquitous environment. *Ubicomp '11 - The 2011 ACM Conference on Ubiquitous Computing, 33–40*. http://doi.org/10.1145/2030092.2030099

Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, S. K., & Chen, J. H. (2009, April). EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine. In *2009 IEEE International Conference on*

Acoustics, Speech and Signal Processing (pp. 489-492). http://doi.org/10.1109/ICASSP.2009.4959627

Lin, Y.-P., Wang, C.-H., Wu, T.-L., Jeng, S.-K., & Chen, J.-H. (2008). Support vector machine for EEG signal classification during listening to emotional music. *2008 IEEE 10th Workshop on Multimedia Signal Processing*, 127–130. http://doi.org/10.1109/MMSP.2008.4665061

Lopes, L. L. (1994). Psychology and economics: Perspectives on risk, cooperation, and the marketplace. *Annual Review of Psychology, 45*, 197-227. http://dx.doi.org/10.1146/annurev.ps.45.020194.001213

Lorist, M. M., Bezdan, E., ten Caat, M., Span, M. M., Roerdink, J. B. T. M., & Maurits, N. M. (2009). The influence of mental fatigue and motivation on neural network dynamics; an EEG coherence study. *Brain Research, 1270*, 95–106. http://doi.org/10.1016/j.brainres.2009.03.015

Lotte, F., van Langhenhove, A., Lamarche, F., Ernest, T., Renard, Y., Arnaldi, B., & Lécuyer, A. (2010). Exploring large virtual environments by thoughts using a brain–computer interface based on motor imagery and high-level commands. *Presence: Teleoperators and Virtual Environments, 19(1),* 54-70. http://doi.org/10.1162/pres.19.1.54

Luay, M., & Revett, K. (2011, July). On the applicability of heart rate for affective gaming. In *WSEAS Special Session on Knowledge Engineering for Decision Support Systems, the CSCC Multiconference, Corfu Island, Greece,* pp. 15-17. ISBN: 978-1-61804-019-0

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288(5472)*, 1835–1838. http://doi.org/10.1126/science.288.5472.1835

MacNamara, A., Ferri, J., & Hajcak, G. (2011). Working memory load reduces the late positive potential and this effect is attenuated with increasing anxiety. *Cognitive, Affective & Behavioral Neuroscience, 11(3)*, 321–331. http://doi.org/10.3758/s13415-011-0036-z

Makeig, S., J. Bell., A., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems, 8*, 145–151. http://doi.org/10.1109/ICOSP.2002.1180091

Makeig, S., Leslie, G., Mullen, T., Sarma, D., Bigdely-Shamlo, N., & Kothe, C. (2011). First demonstration of a musical emotion BCI. In *Affective Computing and Intelligent Interaction* (pp. 487-496). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-24571-8_61

Makeig, S., Leslie, G., Mullen, T., Sarma, D., Bigdely-Shamlo, N., & Kothe, C. (2011). Analyzing brain dynamics of affective engagement. *Lecture Notes in Computer Science, 6975.*

Mason, G., Shuman, T. R., & Cook, K. E. (2013, June). Inverting (flipping) classrooms–Advantages and challenges. In *Proceedings of the 120th ASEE annual conference and exposition, Atlanta.*

MATLAB and Statistics Toolbox Release 2013b. (2013). The MathWorks Inc. Natick, Massachusetts, United States.

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

Mcdowell, K., Lin, C. T., Oie, K. S., Jung, T. P., Gordon, S., Whitaker, K. W.,… & Hairston, W. D. (2013). *Real-World Neuroimaging Technologies. IEEE Access, 1*, 131–149. http://doi.org/10.1109/ACCESS.2013.2260791

McEvoy, L. K., Pellouchoud, E., Smith, M. E., & Gevins, A. (2001). Neurophysiological signals of working memory in normal aging. *Cognitive Brain Research, 11(3)*, 363–376. http://doi.org/10.1016/S0926-6410(01)00009-X

McFarland, D. J., McCane, L. M., David, S. V., & Wolpaw, J. R. (1997). Spatial filter selection for EEG-based communication. *Electroencephalography and Clinical Neurophysiology, 103(3)*, 386–394. http://doi.org/10.1016/S0013-4694(97)00022-2

Mehrabian, A. (1994). *Manual for the revised trait dominance-submissiveness scale (TDS).* University of California.

Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology, 14(4)*, 261–292. http://doi.org/10.1007/BF02686918

Michels, L., Moazami-Goudarzi, M., Jeanmonod, D., & Sarnthein, J. (2008). EEG alpha distinguishes between cuneal and precuneal activation in working memory. *NeuroImage, 40(3)*, 1296–1310. http://doi.org/10.1016/j.neuroimage.2007.12.048

Mikhail, M., El–Ayat, K., Coan, J., & Allen, J. (2013). Using minimal number of electrodes for emotion detection using brain signals produced from a new elicitation technique. *International Journal of Autonomous and Adaptive Communications Systems, 6(1)*, 80-97. http://doi.org/10.1504/IJAACS.2013.050696

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to

complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology, 41(1)*, 49–100. http://doi.org/10.1006/cogp.1999.0734

Mizuhara, H., & Yamaguchi, Y. (2007). Human cortical circuits for central executive function emerge by theta phase synchronization. *NeuroImage, 36(1)*, 232–244. http://doi.org/10.1016/j.neuroimage.2007.02.026

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology, 48(2)*, 229–240. http://doi.org/10.1111/j.1469-8986.2010.01061.x

Mostow, J., Chang, K., & Nelson, J. (2011). Toward exploiting EEG input in a reading tutor. Artificial Intelligence in Education, 230–237. http://doi.org/10.1007/978-3-642-21869-9_31

Mühl, C., Brouwer, A., Wouwe, N. van, van den Broek, E. L., Nijboer, F., & D.K.J., H. (2011). Modality-specific affective responses and their implications for affective BCI. In *Proceedings of the 5th International Brain-Computer Interface Conference* (pp. 4–7). Graz, Austria.

Mühl, C., Heylen, D., & Nijholt, A. (2014). 15 Affective Brain-Computer Interfaces: Neuroscientific Approaches to Affect Detection. *The Oxford Handbook of Affective Computing*.

Mühl, C., Jeunet, C., & Lotte, F. (2014). EEG-based workload estimation across affective contexts. *Frontiers in Neuroscience, 8(June)*, 114. http://doi.org/10.3389/fnins.2014.00114

Müller, M. M., Keil, A., Gruber, T., & Elbert, T. (1999). Processing of affective pictures modulates right-hemispheric gamma band EEG activity. *Clinical Neurophysiology*, 110(11), 1913–20. http://dx.doi.org/10.1016/S1388-2457(99)00151-0

Murray, T. (1999). Authoring Intelligent Tutoring Systems: An analysis of the state of the art. *International Journal on Artificial Intelligence in Education (IJAIED)*, 10, 98–129.

Murugappan, M., Ramachandran, N., & Yaacob, S. (2010). Classification of human emotion from EEG using discrete wavelet transform. *Journal of Biomedical Science and Engineering, 3(4)*, 390–396. http://doi.org/10.4236/jbise.2010.34054

Myrden, A., & Chau, T. (2015). Effects of user mental state on EEG-BCI performance. *Frontiers in Human Neuroscience, 9(June)*, 1–11. http://doi.org/10.3389/fnhum.2015.00308

Niedermeyer, E., Schomer, D., & Lopes da Silva, F. (2011). Niedermeyer's electroencephalography : basic principles, clinical applications, and related fields. Lippincott, Williams and Wilkens.

Nijholt, A., Tan, D., Pfurtscheller, G., Brunner, C., Millán, J. D. R., Allison, B., ... & Müller, K. R. (2008). Brain-computer interfacing for intelligent systems. *IEEE intelligent systems*, *23*(3), 72-79. http://doi.org/10.1109/MIS.2008.41

Nwana, H. S. (1990). Intelligent Tutoring Systems : an overview. *Artificial Intelligence Review, 4*, 251–277. http://doi.org/10.1007/BF00168958

Ochsner, K. N., Hughes, B., Robertson, E. R., Cooper, J. C., & Gabrieli, J. D. E. (2009). Neural systems supporting the control of affective and cognitive conflicts. *Journal of Cognitive Neuroscience, 21(9)*, 1842–1855. http://doi.org/10.1162/jocn.2009.21129

Oliveri, M., Babiloni, C., Filippi, M. M., Caltagirone, C., Babiloni, F., Cicinelli, P., … & Rossini, P. M. (2003). Influence of the supplementary motor area on primary motor cortex excitability during movements triggered by neutral or emotionally unpleasant visual cues. *Experimental Brain Research, 149(2)*, 214–21. http://doi.org/10.1007/s00221-002-1346-8

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: an integrative review of ERP findings. *Biological Psychology, 77(3)*, 247–265. http://doi.org/10.1016/j.biopsycho.2007.11.006

Önal-Hartmann, C., Pauli, P., Ocklenburg, S., & Güntürkün, O. (2012). The motor side of emotions: Investigating the relationship between hemispheres, motor reactions and emotional stimuli. *Psychological Research, 76(3)*, 311–316. http://doi.org/10.1007/s00426-011-0337-4

Onton, J., & Makeig, S. (2009). High-frequency broadband modulations of electroencephalographic spectra. *Frontiers in Human Neuroscience, 3(December)*, 61. http://doi.org/10.3389/neuro.09.061.2009

Onton, J., Westerfield, M., Townsend, J., & Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neuroscience and Biobehavioral Reviews, 30(6)*, 808–22. http://doi.org/10.1016/j.neubiorev.2006.06.007

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 156869. http://doi.org/10.1155/2011/156869

Oostenveld, R., & Oostendorp, T. F. (2002). Validating the boundary element method for forward and inverse EEG computations in the presence of a hole in the skull. *Human Brain Mapping, 17(3)*, 179–192. http://doi.org/10.1002/hbm.10061

Ormrod, J. E. (2011). Human Learning. http://doi.org/10.1017/CBO9781107415324.004

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *The Journal of Neuroscience, 26(28)*, 7523–7531. http://doi.org/10.1523/JNEUROSCI.1948-06.2006

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2003). Response error correction--a demonstration of improved human-machine performance using real-time EEG monitoring. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society, 11(2)*, 173–177. http://doi.org/10.1109/TNSRE.2003.814446

Passarotti, A. M., Sweeney, J. A., & Pavuluri, M. N. (2011). Fronto-limbic dysfunction in mania pre-treatment and persistent amygdala over-activity post-treatment in pediatric bipolar disorder. *Psychopharmacology, 216(4)*, 485–499. http://doi.org/10.1007/s00213-011-2243-2

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. Educational *Psychology Review, 18(4)*, 315–341. http://doi.org/10.1007/s10648-006-9029-9

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. http://doi.org/10.1207/S15326985EP3702

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences, 13(4)*, 160–6. http://doi.org/10.1016/j.tics.2009.01.006

Pfurtscheller, G., & Brunner, C. (2006). Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks. *Neuroimage, 31*, 153–159. http://doi.org/10.1016/j.neuroimage.2005.12.003

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical Neurophysiology, 110(11)*, 1842–1857. http://dx.doi.org/10.1016/S1388-2457(99)00141-8

Pfurtscheller, G., & Neuper, C. (1997). Motor imagery activates primary sensorimotor area in humans. *Neuroscience Letters, 239(2–3)*, 65–68. http://doi.org/10.1016/S0304-3940(97)00889-6

Pfurtscheller, G., & Neuper, C. (2006). Future prospects of ERD/ERS in the context of brain-computer interface (BCI) developments. *Progress in Brain Research, 159*, 433–437. http://dx.doi.org/10.1016/S0079-6123(06)59028-4

Pfurtscheller, G., Neuper, C., Andrew, C., & Edlinger, G. (1997). Foot and hand area mu rhythms. *International Journal of Psychophysiology, 26(1–3)*, 121–135. http://dx.doi.org/10.1016/S0167-8760(97)00760-5

Pfurtscheller, G., Stancák, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band - an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology, 24(1–2)*, 39–46. http://dx.doi.org/10.1016/S0167-8760(96)00066-9

Picard, N., & Strick, P. L. (2003). Activation of the supplementary motor area (SMA) during performance of visually guided movements. *Cerebral Cortex, 13*, 977–986. http://doi.org/10.1093/cercor/13.9.977

Pinegger, A., Wriessnegger, S. C., Faller, J., & Müller-Putz, G. R. (2016). Evaluation of different EEG acquisition systems concerning their suitability for building a brain-computer interface: Case studies. *Frontiers in Neuroscience, 10(September)*, 1–11. http://doi.org/10.3389/fnins.2016.00441

Plass, J. L., Moreno, R., & Brünken, R. (2010). *Cognitive load theory*. Cambridge University Press.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10(2)*, 59–63. http://doi.org/10.1016/j.tics.2005.12.004

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118(10)*, 2128–2148. http://doi.org/10.1016/j.clinph.2007.04.019

Power, M., & Dalgleish, T. (2015). *Cognition and emotion: From order to disorder*. Psychology press.

Raimondo, F., Kamienkowski, J. E., Sigman, M., & Fernandez Slezak, D. (2012). CUDAICA: GPU optimization of Infomax-ICA EEG analysis. *Computational Intelligence and Neuroscience, 2012*, 206972. http://doi.org/10.1155/2012/206972

Rämä, P., Martinkauppi, S., Linnankoski, I., Koivisto, J., Aronen, H. J., & Carlson, S. (2001). Working memory of identification of emotional vocal expressions: an fMRI study. *Neuroimage, 13*(6), 1090-1101. http://doi.org/10.1006/nimg.2001.0777

Ramirez, R., & Vamvakousis, Z. (2012, December). Detecting emotion from EEG signals using the emotive epoc device. In *International Conference on Brain Informatics* (pp. 175-184). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-35139-6_17

Ramoser, H., Müller-Gerking, J., & Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering, 8(4)*, 441–446. http://doi.org/10.1109/86.895946

Reissland, J., & Zander, T. O. (2009). Automated detection of bluffing in a game— Revealing a complex covert user state with a passive BCI. In *Proceedings of the Human Factors and Ergonomics Society Europe Chapter Annual Meeting, Linkoeping, Sweden*.

Reuderink, B., Mühl, C., & Poel, M. (2013). Valence, arousal and dominance in the EEG during game play. *International journal of autonomous and adaptive communications systems*, *6*(1), 45-62. http://doi.org/10.1504/IJAACS.2013.050691

Reuderink, B., Nijholt, A., & Poel, M. (2009, June). Affective Pacman: A frustrating game for brain-computer interface experiments. In *International Conference on Intelligent Technologies for Interactive Entertainment* (pp. 221-227). Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-02315-6_23

Roy, R. N., Bonnet, S., Charbonnier, S., & Campagne, A. (2013). Mental fatigue and working memory load estimation: Interaction and implications for EEG-based passive BCI. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (pp. 6607–6610). http://doi.org/10.1109/EMBC.2013.6611070

Russell, D. M., Moran, T. P., & Jordan, D. S. (1988). The instructional design environment. InJ. Psotka, LD Massey, & SA Mutter (Eds.) Intelligent tutoring systems: Lessons learned (pp. 203-228).

Russell, J. A., & Pratt, G. (1980). A description of the affective quality attributed to environments. *Journal of personality and social psychology*, *38*(2), 311.

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. Journal of Research in Personality, 294, 273–294. http://dx.doi.org/10.1037/0022-3514.38.2.311

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57(3)*, 493-502. http://doi.org/10.1037/0022-3514.57.3.493

Sabatinelli, D., Lang, P. J., Keil, A., & Bradley, M. M. (2007). Emotional perception: Correlation of functional MRI and event-related potentials. *Cerebral Cortex, 17(5)*, 1085–1091. http://doi.org/10.1093/cercor/bhl017

Sanchez, G., Lecaignard, F., Otman, A., Maby, E., & Mattout, J. (2016). Active SAmpling Protocol (ASAP) to optimize individual neurocognitive hypothesis testing: A BCI-inspired dynamic experimental design. *Frontiers in Human Neuroscience, 10(July)*, 1–8. http://doi.org/10.3389/fnhum.2016.00347

Sarlo, M., Buodo, G., Poli, S., & Palomba, D. (2005). Changes in EEG alpha power to different disgust elicitors: The specificity of mutilations. *Neuroscience Letters, 382(3),* 291–296. http://doi.org/10.1016/j.neulet.2005.03.037

Scharinger, C., Kammerer, Y., & Gerjets, P. (2015). Pupil dilation and EEG alpha frequency band power reveal load on executive functions for link-selection processes during text reading. *PloS one*, *10*(6), e0130608. http://dx.doi.org/10.1371/journal.pone.0130608

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology, 52(10),* 1293–1304. http://doi.org/10.1111/psyp.12500

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology, 37(2),* 257–61. http://doi.org/10.1111/1469-8986.3720257

Schutter, D. J. L. G., Putman, P., Hermans, E., & Van Honk, J. (2001). Parietal electroencephalogram beta asymmetry and selective attention to angry facial expressions in healthy human subjects. *Neuroscience Letters, 314(1–2)*, 13–16. http://doi.org/10.1016/S0304-3940(01)02246-7

Shenoy, P., Krauledat, M., Blankertz, B., Rao, R. P. N., & Müller, K.-R. (2006). Towards adaptive classification for BCI. *Journal of Neural Engineering, 3(1),* R13-23. http://doi.org/10.1088/1741-2560/3/1/R02

Shute, V. J., & Psotka, J. (1985). Intelligent tutoring systems: past, present and the future. Science (New York, N.Y.), 228(4698), 456–462. http://doi.org/10.1126/science.228.4698.456

Siemer, M. (2005). Mood-congruent cognitions constitute mood experience. *Emotion, 5(3)*, 296–308. http://doi.org/10.1037/1528-3542.5.3.296

Sigman, M., Peña, M., Goldin, A. P., & Ribeiro, S. (2014). Neuroscience and education: prime time to build the bridge. *Nature Neuroscience, 17(4)*, 497–502. http://doi.org/10.1038/nn.3672

Silberman, E. K., & Weingartner, H. (1986). Hemispheric lateralization of functions related to emotion. Brain and Cognition, 5(3), 322–53. http://dx.doi.org/10.1016/0278-2626(86)90035-7

Singer, E. (2008). Brain Games: Do new controllers that purport to interpret brain activity really work? *Technology Review, 111(4)*, 1–6.

Smedt, T. De, & Menschaert, L. (2012). VALENCE: affective visualisation using EEG. Digital Creativity, (April 2013), 37–41. http://dx.doi.org/10.1080/14626268.2012.719240

Smith, M. E., & Gevins, A. (2005). Neurophysiologic monitoring of mental workload and fatigue during operation of a flight simulator. In *J. A. Caldwell & N. J. Wesensten (Eds.), Defense and Security (pp. 116–126).* http://doi.org/10.1117/12.602181

Spüler, M., Rosenstiel, W., & Bogdan, M. (2012). Principal component based covariate shift adaption to reduce non-stationarity in a MEG-based brain-computer interface. *EURASIP Journal on Advances in Signal Processing, 2012*(1), 1-7. http://doi.org/10.1186/1687-6180-2012-129

Spüler, M., Walter, C., Rosenstiel, W., Gerjets, P., Moeller, K., & Klein, E. (2016). EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning. *ZDM*, 1-12. http://doi.org/10.1007/s11858-015-0754-8

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology (2006), 59(4)*, 745–759. http://doi.org/10.1080/17470210500162854

Steel, P., & König, C. (2006). Integrating theories of motivation. *Academy of Management Review, 31(4)*, 889–913. http://doi.org/10.5465/AMR.2006.22527462

Stengel, E. (1936). Pruefungsangst und pruefungsneurose. Zeitschrift F. Psa. Paed., X, 300–320.

Sweller, J., & Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312. http://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review*, *10*(3), 251-296. http://doi.org/10.1023/A:1022193728205

Szafir, D., & Mutlu, B. (2013). ARTFul: adaptive review technology for flipped learning. *Proceedings of the 2013 ACM Annual Conference on Human Factors in Computing Systems*, 1001–1010. http://doi.org/10.1145/2470654.2466128

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences, 3(4)*, 151–162. http://dx.doi.org/10.1016/S1364-6613(99)01299-1

Tao, J., & Tan, T. (2005, October). Affective computing: A review. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 981-995). Springer Berlin Heidelberg. http://dx.doi.org/10.1007/11573548_125

ten Caat, M., Maurits, N. M., & Roerdink, J. B. T. M. (2008). Data-driven visualization and group analysis of multichannel EEG coherence with functional units. *IEEE Transactions on Visualization and Computer Graphics, 14(4)*, 756–771. http://doi.org/10.1109/TVCG.2008.21

Tomarken, A. J., Davidson, R. J., & Henriques, J. B. (1990). Resting frontal brain asymmetry predicts affective responses to films. *Journal of personality and social psychology, 59*(4), 791. http://dx.doi.org/10.1037/0022-3514.59.4.791

Tomkins, S. S. (2008). Affect Imagery Consciousness: The Complete Edition: Two Volumes. (Vol. 109). Springer Publishing Company. http://doi.org/10.1192/bjp.109.463.837

Torres, F., & Anderson, C. (1985). The normal EEG of the human newborn. *Journal of Clinical Neurophysiology, 2(2)*, 89–103.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46(4)*, 197–221. http://doi.org/10.1080/00461520.2011.611369

Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering, 2*, 157–80. http://doi.org/10.1146/annurev.bb.02.060173.001105

Wager, T. D., Phan, K. L., Liberzon, I., & Taylor, S. F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *NeuroImage, 19(3)*, 513–531. http://doi.org/10.1016/S1053-8119(03)00078-8

Walter, C., Schmidt, S., Rosenstiel, W., Bogdan, M., & Gerjets, P. (2013). Alpha- and theta- frequencies as indicators for optimal cognitive load during learning. In *6th International Cognitive Load Theory Conference* (pp. 3–4). Toulouse, France.

Wang, H., Li, Y., Hu, X., Yang, Y., Meng, Z., & Chang, K. (2011). Using EEG to improve massive open online courses feedback interaction, 59–66. In *AIED Workshops.*

Wang, Guan, C., Phua, K. S., Ang, K. K., Zhang, H., & Kuah, C. (2009). A feasibility study of non-invasive motor-imagery BCI for stroke patients robotic rehabilitation, 10–13. In *2009 4th International IEEE/EMBS Conference on Neural Engineering* (pp. 271-274). http://doi.org/10.1109/NER.2009.5109285

Wass, R., & Golding, C. (2014). Sharpening a tool for teaching: the zone of proximal development. *Teaching in Higher Education, 19(6)*, 671–684. http://doi.org/10.1080/13562517.2014.901958

Watson, R., Latinus, M., Noguchi, T., Garrod, O., Crabbe, F., & Belin, P. (2013). Dissociating task difficulty from incongruence in face-voice emotion integration. *Frontiers in Human Neuroscience, 7(November)*, 1–13. http://doi.org/10.3389/fnhum.2013.00744

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology, 38*, 998–1003. http://doi.org/10.1111/1469-8986.3860998

Weissman, D. H., Gopalakrishnan, A., Hazlett, C. J., & Woldorff, M. G. (2005). Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. *Cerebral Cortex, 15(2)*, 229–237. http://doi.org/10.1093/cercor/bhh125

Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S.,… & Willoughby, J. O. (2007). Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology, 118(8)*, 1877–1888. http://doi.org/10.1016/j.clinph.2007.04.027

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors, 50(3)*, 449–455. http://doi.org/10.1518/001872008X288394

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4(July)*, 1–22. http://doi.org/10.3389/fpsyg.2013.00433

Winkler, I., Jäger, M., Mihajlovic, V., & Tsoneva, T. (2010). Frontal EEG asymmetry based classification of emotional valence using common spatial patterns. *World Academy of Science, Engineering and Technology, 45*, 373-378.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology, 113(6)*, 767–791. http://dx.doi.org/10.1016/S1388-2457(02)00057-3

Wu, L., Eichele, T., & Calhoun, V. D. (2010). Reactivity of hemodynamic responses and functional connectivity to different states of alpha synchrony: A concurrent EEG-fMRI study. *NeuroImage, 52(4)*, 1252–1260. http://doi.org/10.1016/j.neuroimage.2010.05.053

Yeung, A., Garudadri, H., Toen, C. Van, Mercier, P., Balkan, O., Makeig, S., & Virji-Babul, N. (2015). Comparison of foam-based and spring-loaded dry EEG electrodes with wet electrodes in resting and moving conditions. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. http://doi.org/10.1109/EMBC.2015.7320036

Yuan, L., & Powell, S. (2013). MOOCs and open education: Implications for higher education. *Cetis, 19.* http://doi.org/http://publications.cetis.ac.uk/2013/667

Zander, T. O. (2011). Utilizing Brain-Computer Interfaces for Human-Machine Systems. Doctoral thesis. Technische Universität Berlin.

Zander, T. O., Brönstrup, J., & Krol, L. (2014). Towards BCI-based implicit control in human–computer interaction. In *S. H. Fairclough & K. Gilleade (Eds.), Advances in Physiological Computing* (pp. 67–90). London: Springer London. http://doi.org/10.1007/978-1-4471-6392-3

Zander, T. O., Brönstrup, J., Krol, L., Grissmann, S., Gramann, K., & Gerjets, P. (in preparation). Examining increased cognitive load during perceived loss of control.

Zander, T. O., & Jatzev, S. (2009). Detecting affective covert user states with passive brain-computer interfaces. *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, 1–9. http://doi.org/10.1109/ACII.2009.5349456

Zander, T. O., & Jatzev, S. (2012). Context-aware brain-computer interfaces: exploring the information space of user, technical system and environment. *Journal of Neural Engineering, 9(1)*, 16003. http://doi.org/10.1088/1741-2560/9/1/016003

Zander, T. O., & Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering, 8*, 25005. http://doi.org/10.1088/1741-2560/8/2/025005

Zander, T. O., Kothe, C., Jatzev, S., & Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces. In *D. S. Tan & A. Nijholt (Eds.), Brain-Computer Interfaces* (pp. 181–199). London: Springer London. http://doi.org/10.1007/978-1-84996-272-8

Zeidner, M. (1998). Test anxiety: The state of the art. Plenum Press.

Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(1)*, 39–58. http://doi.org/10.1109/TPAMI.2008.52

11 List of Figures