2004

Automatic error recovery using P3 response verification for a brain-computer interface

Samuel A. Inverso

Automatic Error Recovery Using P3 Response Verification for a Brain-Computer Interface

Rochester Institute of Technology
Computer Science Department
Master of Science Thesis

Samuel A. Inverso

May 9, 2004

In partial fulfillment of the requirements for the degree of Master of Science.

Jessica D. Bayliss — May 12, 2004

Advisor: Jessica D. Bayliss, Ph.D.

Michael Van Wie — 12/May/04

Reader: Michael Van Wie, Ph.D.

Roger S. Gaborski — 12 May 2004

Observer: Roger S. Gaborski, Ph.D.

Acknowledgments

First, I would like to thank my advisor, Jessica Bayliss, for the support, freedom, and guidance necessary to succeed in this interdisciplinary field. Jessica's straightforward advice and easy communication style created an environment where the risk inherent in interesting research was positively balanced against the short time frame imposed by a Master's degree.

I would like to thank Gerwin Schalk of the Wadsworth Center for help in understanding BCI2000's subtleties and complexities, and his often immediate responses to email queries. In addition, I would like to thank Agustin Roca and Vijay Kenkat of Tucker-Davis Technologies who were exemplary in their customer service — diagnosing and anticipating hardware and software difficulties with the amplifier, and providing indispensable aid in the signal acquisition system's design.

I would like to thank Carol Marchetti for her elucidation of statistical significance tests, which was instrumental in understanding and applying the correct statistical methods in data analysis.

I would like to thank Rebecca Allen for her allowances and support during my absence from the Media Lab Europe while finishing this thesis.

I would like to thank Andrea Chew, Chad Hulbert, and Juanjo Andres Prado for proof reading this material. Special thanks to Chad for mercilessly scrutinizing this document, which lead to corrections and suggestions that have greatly benefited its presentation.

Thanks to Dan Kunkle for challenging me to achieve and excel throughout this degree.

I would also like to thank all the professors, students, family, and friends who accompanied me throughout this work. They supported me academically, understood my reclusive behavior, and offered little words of encouragement when motivation was minimal. Though their names are not listed, each impacted this thesis in a meaningful way not easily expressed in words.

Finally, I would like to thank my parents, Ronald and Connie, whose early and consistent support of education instilled in me a fortitude without which this work would not have been possible.

Abstract

A brain-computer interface (BCI) is an augmentative communication mechanism that does not rely on peripheral nerves or muscles. Current BCIs are error prone and slow with error rates of 10 to 30% and transmission rates of 10-25 bits/min, however, error recovery and correction in BCI has largely been neglected. The focus of this thesis is the development of a method to automatically recover errors in BCI using the P3 brain signal for response verification.

The existence of the P3 signal in responses to controlled goal items is shown in an evoked potential BCI used to control items in a virtual apartment. A reduced response exists when items are accidentally controlled. Offline experiments were run, and with a theoretical mean improvement in accuracy from 78% to 85%, there was a statistically significant improvement ($P < 0.008$, Wilcoxon signed rank test) in accuracy of 3% using a correlation algorithm for P3 signal detection on responses. The presence of the P3 signal in responses to goal items indicates it can be used for automatic error recovery without requiring additional time, which will improve the speed and accuracy of brain-computer interfaces.

Contents

List of Figures

List of Tables

This page intentionally not left blank.

List of Acronyms

A/D analog-to-digital
ABI Adaptive Brain Interface
AC alternating current
BCI brain-computer interface
CNV contingent negative variation
EEG electroencephalogram
ECG electrocardiogram
EMG electromyogram
EOG electrooculargram
EP evoked potential
ERN error-related negativity
ERD event-related desynchronization
ERS event-related synchronization
ERP event-related potential
fMRI functional magnetic resonance imaging
MEG magnetoencephalography
PET positron emission tomography
RV response verification
SCP slow cortical potential
SEP somatosensory evoked potential
SSAEP steady state auditory evoked potential
SSSEP steady state somatosensory evoked potential
SSVEP steady state visual evoked potential
TTD Thought Translation Device
LSP Language Support Program
VEP visual evoked potential
VR virtual reality
VMRP voluntary motor related potentials

Chapter 1

Introduction

Brain-Computer Interfaces (BCI) are communication and control mechanisms that do not rely on peripheral nerves and muscles [84]. As with all non-perfect information transfer mechanisms BCIs suffer from errors [71]. However, error recovery and correction in BCI has largely been neglected. This thesis describes a method to improve BCI accuracy by automatically recovering errors based on brain signals that occur when the computer provides feedback to a user's selection.

Amyotrophic lateral sclerosis (ALS), multiple sclerosis, cerebral palsy, and spinal cord injury are a few neuromuscular disorders that can completely paralyze individuals but not affect their brains. If restorative treatments are ineffective, the afflicted may live many years with life support systems and twenty-four hour care, but without the ability to communicate or control their environment through normal means. These individuals are locked-in to their bodies [84].

Depending upon the individuals' afflictions, there are three options available for communication and control. First, they may use remaining voluntary muscle control for communication. Some paralyzed people retain control of eye movements, which they can use to answer simple questions or control word processing programs with eye tracking. Second, neural pathways may be reconnected around the break to control healthy muscle. Electromyographic (EMG) signals from muscles unaffected by spinal cord injury can control paralyzed muscles. Finally, if muscular movement is nonexistent, the locked-in individual may employ a direct brain communication mechanism to communicate and control the environment using a computer [83].

The majority of BCI research is rooted in the electrical activity generated by the neurons firing in the brain, electroencephalogram (EEG), because this activity has been throughly studied in cognitive psychology since its discovery by Hans Berger in 1929 [46]. In addition, EEG is fast and cost effective compared to other brain imaging methods (fMRI, SPECT, MEG, etc.); though, it only provides aggregate information on neuronal activity.

Over the past two decades the utility of BCI has been demonstrated using different brain signals, to varying degrees of success, in spelling applications

[24, 9, 53], cursor control [79, 80, 84, 63], virtual environment control [7, 62], games [30, 36], and hand orthosis and direct muscle control by stimulation [55, 61].

Current BCIs are error prone and slow with error rates of 10 to 30% and transmission rates of 10-25 bits/min [57, 83]. Comparatively, a mediocre typist can achieve 750 bits/min (30 words/min). Error recovery and correction in BCI has largely been neglected even though a modest gain from recovery can achieve significant improvements in accuracy and bandwidth. The lack of error recovery methods may be due to BCI's infancy as a field and the neglect of performance errors in cognitive psychology [43], which has resulted in a minimal set of error theories available for application in BCI.

Two types of errors occur in BCI: selecting the incorrect choice and missing the correct choice [53], these are also referred to as false positive and false negative errors respectively. Improved signal recognition has increased BCI accuracy, however, high error rates persist. Some methods of error recovery are: a manual undo choice [52], requiring the user to manually validate the selection with response verification (RV) [82], and detecting the Error-Related Negativity (ERN) [15] or error potential [71] brain signals for automatic response verification.

These techniques have various disadvantages. The manual techniques, undo and RV, slow the overall transmission speed of the BCI and require more choices to be ignored when no errors are present. The automatic techniques, ERN and error potential, do not reduce the speed of the BCI; however, the ERN has only been detected with physical movement (it occurs when the user realizes a mistake was made, such as pressing an incorrect button). In addition, it is unclear what the error potential represents and more experimentation is needed to prove its robustness across individuals and tasks.

For many BCIs, false positive and false negative errors are diametrically opposed because reducing the number of incorrect selections increases the number of incorrect rejections and vice versa. In most cases, it is better to reduce the number of incorrect selections at the expense of incorrect rejections because recovering incorrect selections can be more frustrating as the user must correct the error, while incorrect rejections only require the user to wait for the appropriate choice to be presented again [53]. If automatic error recovery is employed the number of incorrect rejections can be minimized without frustrating the user because incorrect selections are recovered without user involvement.

This thesis presents a new error recovery technique that addresses the disadvantages of current techniques and the difficulty of minimizing these diametric error types. The new technique automatically rejects selections when the P3 brain signal is not present in a response to a selection. Conversely, the the selection is accepted if the P3 signal is present in the response — a response is the feedback from a computer when the user makes a selection, e.g. in a spelling application, if choice 'b' is selected the computer displays the letter 'b' to the user, which is the response to the selection.

In the second chapter the broad background required for brain-computer interfacing is covered. The physical and functional anatomy of the brain is

presented. Different electrical signals beneficial to BCI are overviewed and the BCIs that utilize them are presented. Brain imaging in general is touched upon and EEG is discussed in depth. Finally, BCI errors and performance measurement in bit/rate are explained.

The third chapter is a review of current error recovery strategies. Choice organization's impact on the occurrence of errors is discussed. Manual recovery strategies, undo and manual response verification, along with the automatic recovery strategies, using the brain signals ERN and the error potential, are presented with their advantages and disadvantages in BCI.

The fourth chapter details the automatic error recovery method using the P3 signal for response verification and the results of an offline experiment using this method in a P3 based BCI controlling a virtual apartment. With a theoretical mean improvement in accuracy from 78% to 85%, there was a statistically significant improvement ($P < 0.008$, Wilcoxon signed rank test) in accuracy of 3% using a correlation algorithm for P3 signal detection on responses.

The fifth chapter describes the chess BCI platform for experimentation developed to address the limitations of the virtual reality (VR) apartment offline data analysis study. The platform provides an engaging environment that requires 100% accuracy and is a novel alternative to the traditional BCI speller applications. In addition, a collaboration with researchers at the Institute of Medical Psychology and Behavioral Neurobiology (IMPBN) at the University of Tübingen in Tübingen, Germany to use the chess BCI for one of their patients is reviewed.

The final chapter concludes this work.

This page intentionally not left blank.

Chapter 2

Brain-Computer Interfaces (BCIs)

A brain-computer interface (BCI) is a communication and control mechanism that does not rely on peripheral nerves and muscles [84]. Fundamentally, BCI is a system to record functional activity directly from the brain, recognize the activity recorded, and control a device based on the activity recognized, see Figure 2.1.

Ideally, brain imaging to record functional brain activity should be fast, fine grained, and nonencumbering for effective real-time device control. Practically, brain imaging devices today only partially fulfill these three constraints. Magnetoencephalography (MEG) is fast and fine grained, but requires a room full of equipment and the subject must remain still. Functional magnetic resonance imaging (fMRI) is fine grained, but requires a large machine that cannot be moved, and is slow in recording activity. Electroencephalography (EEG) is fast, inexpensive, and the subject is allowed a limited range of movement, however, it only records aggregate neuronal activity. Among these choices, EEG is the de facto standard in BCI research because it is inexpensive and the activity recorded by EEG is backed by seventy-five years of research experience, while MEG and fMRI are relatively new and expensive technologies.

Recognition of brain activity is limited by the accuracy of the brain imaging technique used and confounded by the cacophony in the brain itself. Many algorithms have been employed to increase the classification rate of EEG BCIs. These methods include simple linear methods, such as averaging EEG signals over multiple trials of the same stimulus [25], as well as correlations between individual trials and predetermined subject averages [6]. Machine learning algorithms have been used to improve BCI accuracy with techniques such as Independent Component Analysis (ICA) for the reduction of artifacts [34] and Support Vector Machines (SVMs) for increasing overall accuracy [42].

BCIs have been used to control various devices such as a hand orthosis (a brace that moves and positions the hand) and an electrical muscle simulator

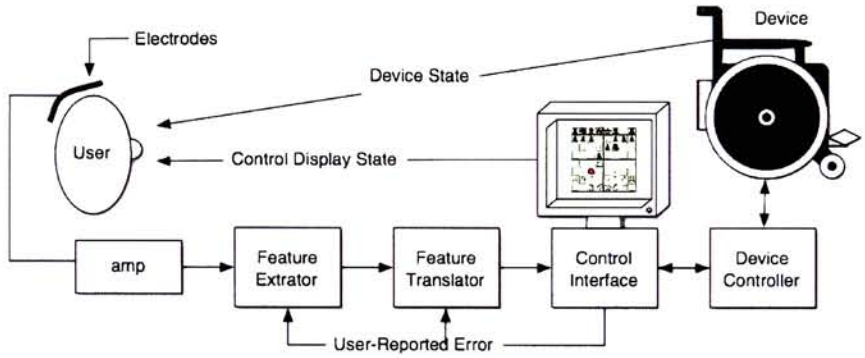


Figure 2.1: General BCI framework proposed by Mason and Birch [40]. Shows the flow of information from the brain to the device and feedback from the device to the person.

[55, 61]. BCI has also been used in computer applications for spelling [24, 9, 53], cursor control [79, 80, 84, 63], games [30, 36], and virtual environment control [7, 62]. Currently BCI is limited by its transmission rate of 10-25 bits/min and error rate of 10-30% [57, 83].

BCI is an interdisciplinary field combining research in cognitive neuroscience, signal processing, and computer science. This chapter covers the broad background required for brain-computer interfacing. The physical and functional anatomy of the brain is described with specific focus on aspects important in EEG BCI. Brain imaging methods are discussed and EEG is detailed in depth including the benefits of different electrical signals and the BCIs that utilize them. Finally, measuring BCI performance in bit/rate is explained.

2.1 The Physical Brain

The adult human brain contains 100 billion neurons spread through the cerebrum, cerebellum, limbic system, and brain stem shown in Figure 2.2 [77]. The most important structure to brain-computer interfaces (BCIs) is the cerebrum's outer four millimeters (mm) of tissue called the cerebral cortex, which contains 20 billion neurons [77] generating the electrical activity that drives the typical electroencephalogram (EEG) BCI.

There are many different types of neurons in the central nervous system (brain and spinal cord) varying in diameter (0.004 mm to 0.1 mm), length (0.15 mm to 2 meters), and shape, among other attributes. The two main neuron classes in the cerebral cortex are pyramidal cells (pyramid-shaped) and stellate cells (star-shaped) [39, 77]. Pyramidal cells are longer than stellate cells and are oriented to form a dipole layer projecting electrical activity to the cortical surface, which is used in an EEG BCI [26]. Neurons consist of three main structures: the soma, dendrites, and axon depicted in Figure 2.3.

The soma is the neuron's cell body and contains typical cell structures such

Figure 2.2: Brain Structures: cerebrum, cerebellum, limbic system, and brain stem. Used with permission [3].

as a nucleus, mitochondria, and ribosomes. Dendrites branch from the soma forming a tree shape. The dendrites are the inputs into the neuron and are covered with many synapses, which are the input points. The axon is the neuron's output. Unlike the branches formed by dendrites, the axon has a single connection to the soma called the axon hillock. The axon resembles a long thread with frayed ends. The ends of the axon form synapses with dendrites where information is transfered [39].

The dendrites and axons of different neurons stop short of touching each other at the synapse. Instead, a gap is formed called the synaptic cleft and the electrical impulse that travels through the axon is converted to a chemical signal that traverses the gap and is converted back to an electrical impulse by the receiving dendrite [39].

Neurons are the brain's decision makers, forming intricate webs with each other giving rise to the brains processing power. Each neuron may have 1,000 to 10,000 connections with neurons in their immediate area or far parts of the brain [4, 77]. When a neuron receives input an electrical potential is created, when this potential reaches a threshold the neuron 'fires', sending an electrical impulse down its axon towards other neurons. A neuron can excite other neurons to fire or inhibit other neurons from firing using this process. [39, 45]

Neuroscientists have shown that groups of neurons will fire in the same location on the cerebral cortex based on function. This has given rise to the functional map of the brain shown in Figure 2.4. In the functional map, the brain is divided into four lobes: frontal, parietal, occipital, and temporal [39].

The *frontal* lobe extends from behind the eyes to the top of the head and is responsible for analysis, planning, decisions, movement and motor skills, language production, and emotions. The *parietal* lobe extends from the frontal lobe

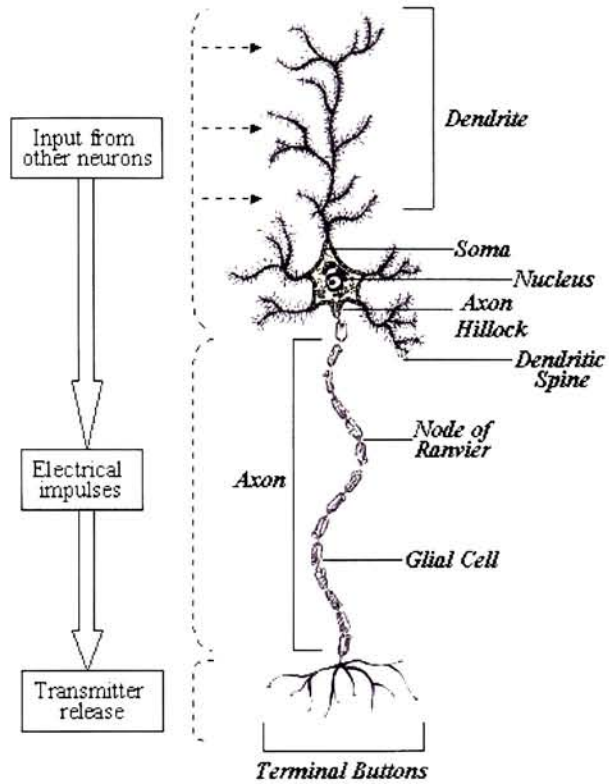


Figure 2.3: Projection neuron showing information flow. Used with permission [27].

to about where the skull begins to steeply slope downwards and is responsible for body location and is the receiver of sensory information from the body. The *occipital* lobe extends from the parietal lobe to the back of the head and sits above the cerebellum. The occipital lobe is responsible for processing visual information and has a direct link with the eyes. Finally, the *temporal* lobe extends along both sides of the head parallel to the ears and touches all three of the other lobes. The temporal lobe is involved with speech comprehension, recognizing objects, scenes, and faces, and maintaining autobiographical information in conjunction with the frontal lobe [65, 37].

2.2 Measuring Brain Activity

There are a variety of imaging devices that can be used for BCI. These devices include electroencephalogram (EEG), magnetoencephalography (MEG), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI) described in Table 2.1.

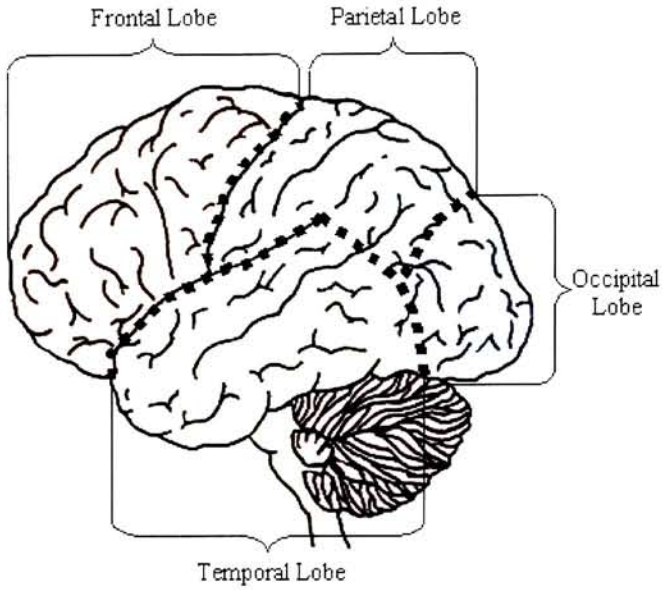


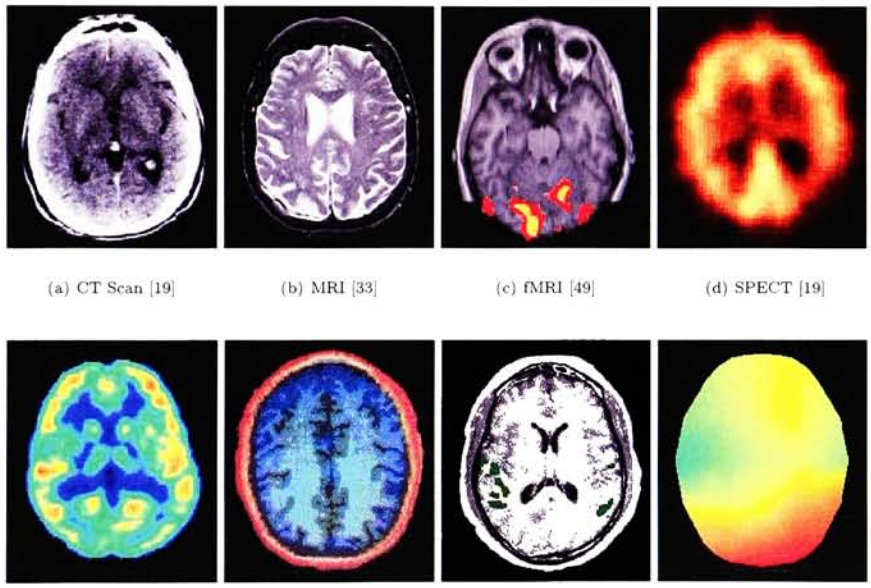
Figure 2.4: The brain's functional lobes. Used with permission [27].

Table 2.1: Brain imaging methods that can be used for BCI [27].

Method	Description
Electroencephalography (EEG)	Maps general brain activity using scalp electrodes.
Magnetoencephalography (MEG)	Measures magnetic fields generated by electrical currents at cell level.
Positron Emission Tomography (PET)	The subject ingests radioactive tagged glucose. After the glucose enters the blood stream the PET machine measures the concentrations of glucose, which corresponds to the brain's active areas.
Single-Photon Emission Computed Tomography (SPECT)	Similar to PET, but with poorer spatial resolution because it only measures a single photon.

Method	Description
Functional Magnetic Resonance Imaging (fMRI)	Based on Magnetic Resonance Imaging (MRI) technology. fMRI can detect oxygen levels in blood to show variations in neural activity without ingesting radioactive markers.

Figure 2.5 depicts spatial resolutions for different brain imaging techniques. Computer Tomography (CT) scans are the most fined grained, approximately 0.3 mm to 1 mm. Unfortunately, CT uses X-rays to scan the brain, which can damage cells. In addition, it does not provide functional information about the brain's activity. EEG provides a good functional map of the overall brain activity but has poor spatial resolution, ranging from 26.6 mm to 35.3 mm. Neurons' widths range from 0.004 mm (granule cell) to 0.1 mm (motor neuron) [77].



(a) CT Scan [19] (b) MRI [33] (c) fMRI [49] (d) SPECT [19]

(e) PET [73] (f) PET/MRI (Multi-Modal) [81] (g) MEG (Superimposed on MRI) [27] (h) EEG [68]

Figure 2.5: Example brain scans depicting spatial resolution. Computer Tomography (CT) is an X-ray imaging technique that does not provide functional information. Used with permission.

Figure 2.6 compares imaging methods' spatial resolution to temporal resolution. MEG provides the best spatial and temporal resolution, however, it

requires a room full of expensive monitoring equipment and the subject's head cannot move during the recording. While EEG spatial resolution is poor, it has very good temporal resolution, which is important when translating rapidly changing brain activity to users' desires. EEG signals are widely used in BCI research because they are fast temporally, inexpensive and less cumbersome than other methods (MEG, fMRI, SPECT, and PET require a room full of equipment), and non-invasive (the user wears a cap of electrodes) [27]. In addition, EEG signals can easily be combined with other techniques, such as fMRI, for BCI systems that require fine grained spatial resolution. For these reasons most BCI research, including this thesis, focuses on EEG.

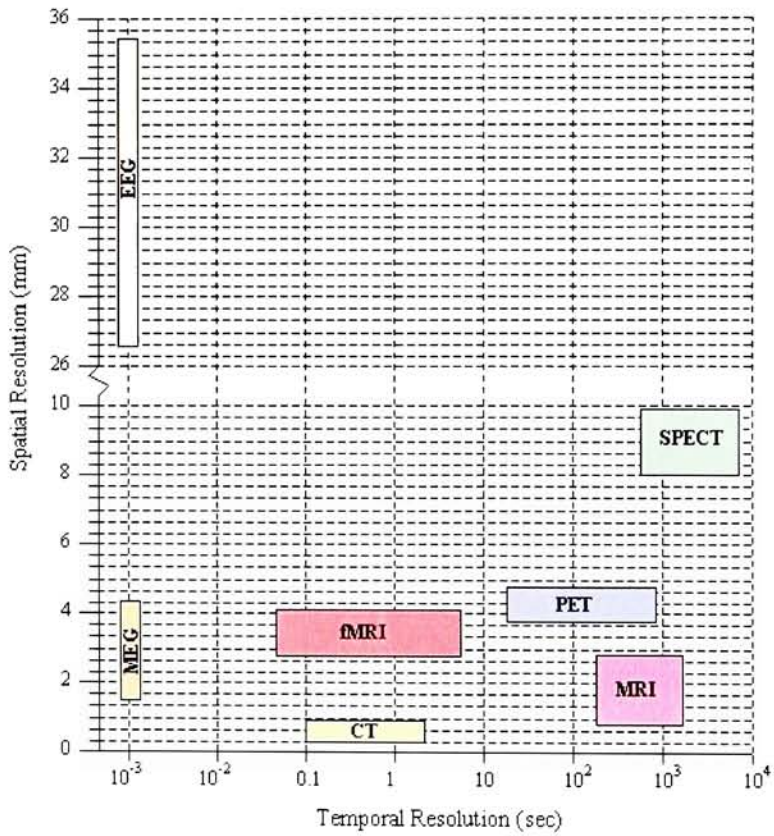


Figure 2.6: Spatial (mm) vs Temporal (sec) Resolutions for Brain Graphing Methods. Used with permission [27].

2.3 Recording EEG Activity

This section details recording EEG from the scalp, known as surface EEG. To record EEG four types of equipment are required: electrodes, an operational-amplifier, an analog-to-digital (A/D) converter, and a computer. The electrodes acquire analog electrical signals from the scalp, which are then sent to the operational-amplifier for amplification. After amplification the signal is digitized by the A/D converter and transfered to a computer where they are interpreted in real-time using signal processing algorithms or stored for later processing. In BCI literature, interpreting signals for communication and control while the subject is using the BCI is referred to as *online*, while processing stored signals is referred to as *offline*.

As described in Section 2.1, when neurons fire they generate electrical activity that summates in the scalp. This electrical activity creates different electrical voltages (potentials) on the scalp, which are detected by electrodes. The voltages on the scalp from the neuron firing is very small, typically at most 50 μV, and needs to be increased to the sensitivity range of the A/D converter, which is usually ± 0.5 V to ± 10 V before further processing.

The scalp voltage is increased to the A/D converter's range with an operational amplifier, which is similar to an audio amplifier found in car and home theater systems. Amplifiers increase the voltage by a multiplicative constant called *gain*, as shown in Equation 2.1 where V_i is the input voltage, g is the gain, and V_o is the amplified output voltage.

$$V_o = V_i \times g \tag{2.1}$$

For example, a typical EEG amplifier gain is 50,000. Given an input of 10 μV the output will be 500,000 μV, which is 0.5 V, and within the range of an A/D converter.

The electrical potentials on the scalp are analog signals that need to be converted to digital representations for a computer to process them. Converting an analog signal to digital form is called *sampling*, because the continuous analog signal is sampled (recorded) at discrete time intervals. Two important aspects of analog to digital conversion are the sampling rate and resolution.

The *sampling rate* indicates how often the A/D converter samples the continuous signal. This rate is represented in units of Hertz (Hz), which are 1/seconds. If the sampling rate is 128 Hz, then the signal is sampled 128 times a second.

The faster the sampling rate the more accurate the signal will be represented in digital form. The Nyquist theorem states the highest frequency that can be accurately reconstructed without error is half the sampling rate. In other words, to represent an analog signal without error in digital form the signal must be sampled at least twice its frequency [26]. As shown in Section 2.4, brain signals have frequencies up to 30 Hz. To accurately represent 30 Hz beta signals in digital form the Nyquist theorem states the signal must be sampled at least at 60 Hz (2×30 Hz). If the beta signal is sampled at less then 60 Hz false frequency components will appear in the reconstructed signal. The false frequency error resulting from sampling too slow is called *aliasing*.

The A/D converter's *resolution* also contributes to its accuracy in representing the signal in digital form. Resolution is given in bits — a bit is the fundamental unit of binary computing and can either be a 0 or 1. A resolution of 12 bits represents a signal using 4,096 points ($2^{12} = 4{,}096$), a resolution of 16 bits has 65,536 points, and a resolution of 2 bits has 4 points. The resolution points are evenly spread across the A/D's voltage output range and indicate the precision of the signal's recording, which is a practical consideration when comparing signals.

For example, an A/D converter with 12 bit resolution and output range 5 V to +5 V measures the signal in 0.002 Volt increments (10 V total range, 4,096 points, 10 V/4096 = 0.002 V) An A/D converter with 16 bit resolution and output range -5 V to +5 V measures the signal in 0.0001 Volt increments (10 V/65536 = 0.0001 V). The choice of resolution is based on the fidelity the BCI signal processing requires.

2.3.1 Electrodes

Electrodes conduct the electrocortical potentials from the scalp to an amplifier. Surface EEG electrodes are 8 to 9 mm in diameter and are composed of silver-silver chloride (Ag/AgCl), gold, or tin [26]. Needle electrodes are generally not used for BCI as they are uncomfortable for the subject, may cause infection, and have poor recording quality [66]. Throughout this thesis 'electrode' refers to surface electrode unless otherwise stated.

Electrodes are cup-shaped to hold electrolytic paste or gel depending on how they are affixed to the scalp. The electrolytic paste and gel aid the conductivity of the electrodes and help prevent motion artifact. For good quality recordings, the electrodes must be firmly attached to the scalp and the electrode impedance should between 100 and 5,000 Ω (Ohms) [26].

Impedance is the resistance to current flow. If the impedance between the electrode and scalp is high, the brain's electrical activity will not be conducted through the electrodes properly and large differences in impedances between electrodes favors 60 Hz noise [26], see Section 2.3.5. Electrode impedance is measured by sending a weak alternating current through one electrode and recording it from a second electrode. With proper electrode application impedances can be reduced to less than 3,000 Ω [64].

Only meters specifically designed to test scalp electrode impedances should be used. Impedance meters designed to test electrical circuits may send a large painful current to the patient, and can also polarize the paste or gel changing their conductive properties [26].

2.3.2 Electrode placement

The International 10–20 System of Electrode Placement [32], seen in Figure 2.7(a), was developed in 1958 by electroencephalographers to compare data using a common terminology and reference system, and defines the placement of twenty-one electrodes on the scalp. The system's name derives from the equal spacing

of electrodes 10% and 20% of the distance between skull landmarks [29].

The electrode position names consist of a letter followed by a number. The letter indicates the structural lobe below the electrode: **F**rontal, **C**entral, **P**arietal, **O**ccipital, and **T**emporal. Numbers indicate if the electrode is left, right, or on the midline. Odd numbers indicate left, even numbers indicate right, and 'z' for 0 indicates on the midline. The system uses every other odd and even number starting with three and four to leave room for additional electrode placements.

For example, from Figure 2.7(a), O1 is on the left side of the head over the occipital lobe, and Cz is the position at the very top of the head over the central lobe on the medial line of the left and right cerebral hemispheres.

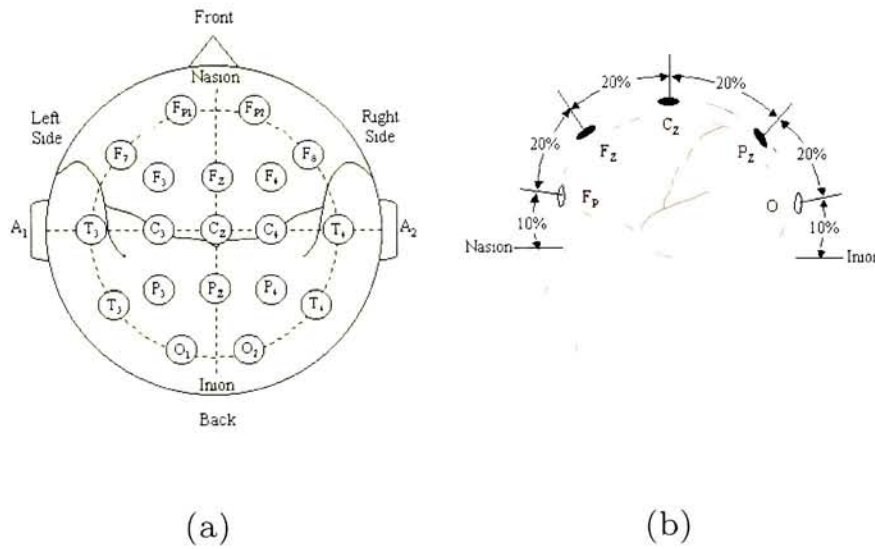


Figure 2.7: International 10-20 System of Electrode Placement. (a) Top view showing the 21 electrode positions. (b) Side view showing electrode positions at 10 and 20% of the distance between the nasion and inion measured over the top of the head. Used with permission [27].

2.3.3 Reference and Bipolar Recordings

EEG measures the potential difference (voltage) between two electrodes, i.e. the electrical signal from one electrode is subtracted from the other. In a bipolar recording, the two electrodes both measure cortical activity. In a reference recording, one electrode measures cortical activity while the other is placed on a body part without cortical activity, such as an earlobe. When recording in

reference all electrodes measuring cortical activity are linked to a single reference electrode [64].

2.3.4 Grounding

Equipment and subjects must be properly grounded during EEG recordings for safety and to reduce noise [64]. A bony protuberance, such as the mastoid behind the ear, should be connected with an electrode to the EEG amplifiers electrode ground jack.

2.3.5 Artifacts

Artifacts are non-cerebral activity in the EEG that may masquerade as brain activity and otherwise obscure real brain activity. Artifacts appear because of the large amplification required to measure electrical brain activity, and may be technological or physiological in nature. The following is a list of the major artifacts that may occur when performing BCI experiments [66].

Technical artifacts

- Electrode Improperly attached electrodes and faulty wires can cause artifacts. An impedance checker should be used to verify that electrodes are properly attached. Electrode impedances should be less than 5,000 Ω [64].

  Faulty and partially broken wires are difficult to diagnose as they may intermittently cause artifacts. Faulty wires may be identified by looking at the activity in nearby electrodes; because EEG measures the potential field across the head, electrodes near each other should show similar activity, dissimilar activity may be a result of technical artifacts [64].

- 60 Hz interference Nearby electrical equipment and power lines can induce rhythmic 60 Hz cycles (50 Hz in Europe) from alternating current (AC) power. Proper grounding of the subject and equipment can reduce this interference. In addition, amplifiers typically have 60/50 Hz notch filters that remove this noise before amplification [64].

- Ground loop - 60 Hz interference that occurs when a ground electrode is shorted to an active electrode [64].

Physiological artifacts

- Oculographic - Eye movement and blinking can generate spikes of 200 to 400 μV, which travel posteriorly. Eye electrical activity recordings, electrooculargram (EOG), are used to differentiate eye activity from cerebral activity. EOG is typically recorded from electrodes above and below the center of the eye, or above the eye and on the outer canthus [66]. In BCI,

experimenters will generally reject trials with amplitudes greater than 50–90 μV when classifying activity to account for electrooculargram (EOG) artifact. Alternatively, regression algorithms may be used to remove eye artifacts from trials that would otherwise be rejected [72].

- Muscle - Muscle contractions, such as talking, jaw clenching, and chewing, generate electrical activity, which may be seen in EEG [26]. Electrical muscle activity, electromyogram (EMG), is dominant in the 50 to 150 Hz frequency range and peaks at 5 mV [38, 18]. A lowpass filter is used to reduce muscle artifacts if the cerebral activity of interest is low enough.

- Cardiac Electrical activity from the heart, electrocardiogram (ECG), can appear in EEG. Pulse waves caused by electrodes over blood vessels and ballistocardiographic artifact caused by a subject's rocking movement when the heart beats may appear in the EEG [64].

- Perspiration Salt in perspiration creates a "salt bridge" between electrodes that decreases inter-electrode impedances causing shorts. Perspiration may also separate the electrode from the scalp overtime [66, 26].

- Motion Head and body movement may move the leads and contacts, which causes wide wave artifacts. This can be corrected by asking the subject not to move [26, 66].

2.4 Electroencephalogram (EEG)

Electroencephalogram (EEG) is the recording of electrical activity from the brain, discovered in humans by Hans Berger (1873–1941) in 1929 [46]. The electrical activity generated from neurons firing, Section 2.1, summates on the human scalp. Using EEG recording equipment, Section 2.3, this electrical activity can be recognized and utilized in brain-computer interfaces (BCIs). This section reviews the signals generated by the brain, where they are recorded on the scalp, their functional significance, and their use in BCI.

2.4.1 Rhythmic Brain Activity

Rhythmic brain waves were the first brain signals discovered by Hans Berger [46]. Rhythmic brain waves are brain signals that occur continuously and repeat in amplitude, frequency, and waveform [26]. These rhythmic signals are divided into frequency ranges named after Greek letters: delta δ, theta θ, alpha α, and beta β, see Figure 2.8. The mapping of Greek letters and frequency ranges follows their chronological discovery, not a logical increasing in frequency.

Alpha

The alpha rhythm ranges between 8 and 13 Hz in normal adults and is distributed over the occipital, posterior temporal, and parietal areas. The wave-

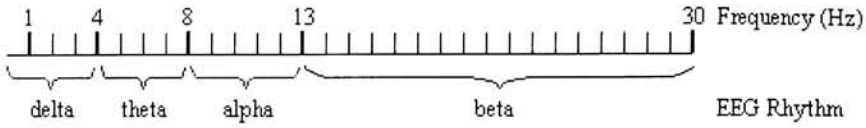


Figure 2.8: Delta, Theta, Alpha, and Beta rhythmic EEG frequency ranges. Used with permission [27].

forms are monomorphic (regular in shape) with sharp points at the top or bottom, or sinusoidal. Alpha's amplitude is variable, but averages 50 μV [26].

Alpha activity is typically constant in its frequency in individual subjects, but varies across subjects, and may decrease by 1 Hz or more with drowsiness. It increases for a short duration after eye closure, and is blocked by eye opening, i.e. its amplitude significantly decreases. Its frequency should be the same in both hemispheres at any given time, while the amplitude may differ between the two hemispheres [26].

The alpha rhythm's purpose is unknown; though, the posterior distribution, eye opening and closing influence on the wave, and other characteristics indicate it is "integrated with the visual system function and possibly represents activity which appears in absence of specific input to that system [26]."

The rhythm is named 'alpha' for purely historical reasons. When Hans Burger made his measurements he named the first rhythm he identified 'alpha' after the first letter in the Greek alphabet: α [46].

Beta

The beta rhythm's frequency band is between 13 and 30 Hz and appears in three main types — frontal, widespread, and posterior — which vary in distribution and reactivity. These three beta patterns disappear in sleep, but frontal and widespread activity remain longer than alpha activity during drowsiness when beta becomes more dominant. Beta rhythm amplitude is typically lower than alpha amplitude, seldom exceeding 30 μV [26]. Beta activity greater than 30 Hz is often referred to as gamma activity.

The most common type of beta rhythm is distributed frontally and extends into the central regions. This beta rhythm may be blocked by movement, intention to move, and tactile stimulation [26].

Widespread beta rhythm appears over the majority of the head and is not blocked by any stimulus [26].

Posterior beta rhythm is also known as fast alpha variant because it is located in the same area as the alpha rhythm, is blocked the same as alpha rhythms, and intermixes, alternates, and replaces alpha rhythms. In addition, its frequency is usually twice alpha's frequency (16–20 Hz) [26].

The beta rhythm's purpose is unknown, but the "blocking mechanism of frontal beta suggests a relationship between this type of beta rhythm and sensorimotor functions of the underlying cortex [26]."

Mu

The mu rhythm's frequency band is between 7 and 11 Hz and appears in the central and centroparietal regions as arch-shaped trains lasting a few seconds. Mu rhythms are more apparent during alpha blocking because mu and alpha rhythms overlap in frequency range [26].

Similar to beta rhythms, the mu rhythm is blocked by both voluntary and involuntary movement, intention to move, and tactile stimulation [26].

The mu rhythm's purpose is unclear, but "may be related to somatosensory processes associated with movement [26]." The blocking effect caused by movement and tactile stimulation may indicate the mu rhythm represents "the idling of a sensory system not processing specific input from the thalamic nuclei [26]."

Delta and Theta

The delta rhythm's frequency band is below 4 Hz; theta's is between 4 and 8 Hz. Both rhythms primarily occur in deep sleep [26] making them of little use in brain-computer interfaces; though, theta rhythms do appear in small unorganized amounts in normal adults [26, 37].

BCIs based on Rhythmic Activity

Wolpaw et al. [84] developed a BCI using self-regulation of mu and beta rhythm amplitude. After 5–10 half hour training sessions subjects can learn to increase and decrease the amplitude of the mu or beta rhythm to control a mouse cursor, and with sufficient training can achieve information transfer rates up to 25 bits/min.

Pineda et al. [62] developed a BCI based on the difference of mu power at C3 and C4 using Cz as ground. Users were trained for 10 hours over five weeks to generate mu activity under both C3 and C4 to move left in a virtual environment, and to generate less mu in one hemisphere than the other to move the virtual character right. Transfer rates were not reported, but all subjects were able to demonstrate both types of control.

Brainball is a novel game developed by Hjelm and Browall [30] where a player attempts to roll a ball into an opponent's goal by achieving a higher state of relaxation measured as the ratio of alpha to beta waves. Alpha and beta self-regulations are rarely used in BCIs because they require shifting between states of relaxation and alertness. These waves have traditionally been restricted to neurofeedback/biofeedback applications [1].

Doherty et al. [22, 21, 23] studied the brain-body interface Cyberlink [2], which uses a combination of electromyogram (EMG), EOG, and alpha and beta EEG for controlling various interfaces. They reported subjects with traumatic brain injuries achieved accuracy rates between 44 and 100%. Though it is unclear to what degree Cyberlink's control is directly from the brain.

Guger et al. [28] reported a large scale imagined motor movement BCI experiment with ninety-nine healthy subjects. After 20–30 minutes of training, 93% of the subjects were able to achieve 60% accuracy by imagining moving

their right hand (which extended an on screen bar right) or both feet (which extended an onscreen bar left).

2.4.2 Event Related Potential (ERP)

Event related potentials (ERP) are changes in the brain's electrical activity in response to stimuli. Unlike rhythmic brain activity, ERPs are short signals with definitive beginnings and endings. These signals are also time-locked to a stimulus, which can be auditory, visual, or somatosensory (touch). The following is not an exhaustive list of ERPs, but a list of ERPs that have utility in BCI.

P300

The P300 is a large brain signal that is evoked by novel and task relevant stimuli. These forms are designated P3a (novelty) and P3b (task relevance). The P300's name reflects that it is a positive signal that peaks **300** ms after the onset of stimulus. Most time related evoked potentials are named in this manner [45]. The P300 is also referred to as the P3.

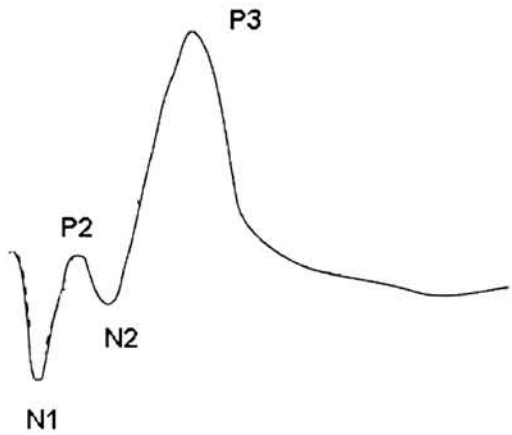


Figure 2.9: P300 waveform.

The P3a appears frontally when the subject is aware of a novel or infrequent stimulus. For example, if the subject is presented with many low frequency tones intermixed with infrequent high frequency tones (e.g. lows tones occur five times as often as high tones), the high tones will evoke the P300 [45]. P3a experiments are often called *oddball* experiments referring to the infrequent stimuli being rare and different from frequent stimuli.

The P3b appears parietally when the subject is aware of a task relevant stimulus. For example, if the subject is repeatedly presented letters of the alphabet one letter at a time and is given a spelling task, the P300 will occur when the current letter the subject wants to 'type' is presented [45].

Bennington [8] demonstrated the P300's amplitude is smaller when the subject is passively aware of the stimulus vs actively attending to the stimulus (such as counting when the stimulus appears) in oddball auditory and visual experiments. The P300's amplitude is also smaller and may disappear if the subject is drowsy. The latency of the P300 increases with older subjects [45].

The source site for the P300 in the brain is unknown; however, it is thought to be generated in the middle of the temporal lobe (medial-temporal); though, this is unproven. Because the P300 is evoked by many different tasks there may not be one source generator [45].

Steady State Visual Evoked Potential (SSVEP)

Steady state visual evoked potentials (SSVEP) are rhythmic waves that occur occipitally when the subject is fixating on a flashing stimulus. The waves have the same frequency as the repetition rate of the flashing stimulus. SSVEPs occur when the stimulus flashes 10 times per second and up to 62 times per second. Flashing beyond 62 times per second exceeds the average critical frequency of photic driving (CFPD) [45].

Contingent Negative Variation (CNV)

The contingent negative variation (CNV) signal occurs when a subject is required to respond to an imperative stimulus that occurs following a warning stimulus, i.e. it occurs for expectation [45].

Typically, the warning and imperative stimulus are separated by 1–2 seconds, and the contingent negative variation (CNV) begins 400 ms after the warning stimulus, peaking up to -50 μV at 800 ms [45].

The subject must expect and intend to respond to the imperative stimulus for the CNV to occur [45].

BCIs based on ERPs

The P300's robustness across multiple individuals and the lack of training required to utilize it makes it a desirable signal for BCIs. Donchin et al. [24] used the P300 in a spelling BCI. Subjects were presented a six-by-six matrix of letters and numbers with intensifying rows or columns every 125 ms as stimuli, see Figure 2.10. The subjects concentrated on a letter and mentally counted every time it intensified. In online trials 56% accuracy was achieved; offline processing showed communication rates of 7.8 characters/min at 80% accuracy and 4.8 characters/min at 90% accuracy. Bayliss [7] explored the P300 in environmental control using a virtual-reality environment where subjects were able to turn on or off a light, tv, and radio, and were able to make a virtual figure appear and disappear with Hi and Bye commands. Online accuracies ranged between 67% and 80% across nine able-bodied subjects. Polikoff et al. [63] used the P300 for cursor control using four compass points as stimuli.

One of the first BCIs was created in the 1970s by Jacque Vidal [79, 80]. Vidal used the visual evoked potential (VEP) recorded over the scalp to move a cursor

Figure 2.10: Farewell and Donchin's P300 speller interface. Rows and columns intensify ever 125 ms. Image shows row intensification [24].

based on where the subject gazed. Sutter (as cited in [83]) created a similar VEP system for spelling that used an 8×8 grid of letters that flashed red/green 40–70 times/second in overlapping subgroups. In Sutter's experiments, healthy subjects were able to spell 10–12 words/min. Middendorf et al. [44] used a different VEP approach by flashing multiple buttons simultaneously at different rates. The button the subject gazed at was determined by the frequency of the photic drive response over the visual cortex. VEP BCIs are considered 'dependent' BCIs because they depend on the ability of the user to control the direction of eye gaze.

2.4.3 Event-Related Desynchronization (ERD) and Event-Related Synchronization (ERS)

Lehtonen [37] succinctly defines event-related desynchronization (ERD) and event-related synchronization (ERS) as follows:

Event-Related Desynchronization (ERD) : is an amplitude *attenuation* of a certain EEG rhythm.

Event-Related Synchronization (ERS) : is an amplitude *enhancement* of a certain EEG rhythm.

Event-Related Desynchronization (ERD) and Event-Related Synchronization (ERS) are measured as a percent of increase and decrease in a frequency band's power (amplitude) relative to a reference interval. The reference interval is a block of EEG that occurs well before the event, usually during a resting phase, and should not contain EEG associated with the onset of the event [60].

A pattern of ERD followed by ERS occurs in motor movement over the motor cortex. For example, two seconds prior to moving a finger ERD occurs as

the attenuation of the alpha band (8–13 Hz) in the motor cortex, predominantly in the hemisphere contralateral to the moving finger. ERS occurs at the onset of finger movement in the low beta band (16–21 Hz) in the same area. Similar ERS following ERD patterns are also seen in foot and hand movement; though, in slightly shifted beta bands [60].

BCIs based on ERD and ERS

Pfurtscheller et al. [58, 59] showed event-related desynchronization (ERD) and event-related synchronization (ERS) appear with imagined motor movement, and have extensively explored their use in the Graz BCI [56, 54, 57]. In offline evaluation, subjects achieved 80–100% accuracy with information transfer rates of 3 to 17.2 bits/min [57]. In addition, Pfurtscheller's group has used ERD and ERS to control a hand orthosis and for direct stimulation of muscles to restore grasp function in tetrapalegics [55, 61].

2.4.4 Other BCI approaches using EEG

Slow Cortical Potential (SCP)

Slow cortical potentials (SCPs) are slow voltage changes generated in the cortex that occur over 0.5 to 10.0 seconds [83]. Birbaumer et al. [9, 11, 10] studied users' ability to control the SCP's positive and negative amplitude with applications in BCI. Birbaumer's group developed the *Thought Translation Device* (TTD), which uses slow cortical potentials (SCPs) for two choice selection. In the TTD, a cursor moves from left to right across a screen at constant velocity. The cursor moves towards the top choice as the SCP becomes positive and the cursor moves towards the bottom choice as the SCP becomes negative. Selecting a choice takes 4 seconds. Users train in 1-2 hour sessions/week over weeks or months until they consistently achieve 75% or better accuracies, then they are switched to the *Language Support Program* (LSP).

The LSP [53, 52] extends the TTD two choice selection technique to spelling by presenting the user with two banks of letters. Each letter bank contains half of the alphabet. When the user selects a bank it is split in half and presented to the user again to chose. In this way the user can progressively split the bank in half to reach the desired letter. Users have achieved accuracies between 65 and 80% with information transfer rates of 0.75 to 15 bits/min.

Voluntary Movement Related Potential (VMRP)

Voluntary Movement Related Potentials (VMRPs) occur over the sensorimotor cortex during actual or imagined movement, e.g. moving a finger. Birch et al. [12, 13] demonstrated the use of VMRP recorded with "six bipolar signals over the supplementary motor area (SMA) and sensory-motor cortex ($F_1 - FC_1, F_z - FC_z, F_2 - FC_2, FC_1 - C_1, FC_z - C_z$, and $FC_2 - C_2$)" in the 1–4 Hz band. In a study with four able-bodied and three spinal cord injured subjects their BCI

was able to classify the presence of VMRP elicited by imagined motor movement with greater than 94% accuracy [13].

Mental Tasks

Spontaneous EEG, resulting from a mental task, can also be to drive a BCI. Suppes [74] performed a study using EEG to discriminate seven mentally spoken words (first, second, third, yes, no, left, and right) with 71 to 86% accuracy. Ryu et al. [67] achieved an 80% success rate discriminating between mental 'Yes' and 'No' with twelve subjects using artificial training neural networks. These studies did not attempt to control online BCI; however, Millan et al. [20] reported the Adaptive Brain Interface (ABI) uses a neural classifier to discriminate between mental states with online results.

In the Adaptive Brain Interface (ABI) [20], subjects choose three of five mental tasks (relaxation, imagined left and right hand or arm movement, mental cube rotation, subtraction, and word association). A classifier is trained over four sessions at five minutes a session with the subject performing the mental task for 10 to 15 seconds. The classifier determines the subject's mental state every 0.5 seconds with a theoretical maximum information transfer rate of 2 bits/minute. In real usage, subjects were able to achieve 0.22 to 0.68 bits/minute.

2.5 Errors in a BCI

Two types of errors occur in BCI: selecting the incorrect choice and missing the correct choice [53], these are also referred to as false positive and false negative errors respectively.

For many BCIs, the minimization of false positive and false negative errors are diametrically opposed because reducing the number of incorrect selections increases the number of incorrect rejections and vice versa. In this case, it is better to reduce the number of incorrect selections at the expense of incorrect rejections because recovering incorrect selections can be more frustrating as the user must correct the error while incorrect rejections only require the user to wait for the appropriate choice to be presented again [53]. If automatic error recovery is employed the number of incorrect rejections can be minimized without frustrating the user because incorrect selections are recovered without user involvement.

2.6 BCI Performance Measures

Wolpaw et al. [82] proposed measuring BCI performance in bit rate to facilitate the comparison of BCIs in different domains, such as spelling and cursor movement. In order to calculate the bit rate, the bits/trial must first be found

with the equation [82]:

$$B_t = \log_2 N + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{N - 1} \tag{2.2}$$

where there are N possible selections, each of which are equally probable, and P is the accuracy of selecting without error. The bit rate B_m, is the bits/trial B_t, multiplied by the average number of trials/minute T_m:

$$B_m = B_t T_m \tag{2.3}$$

Chapter 3

Related Work

The focus of this thesis is the development of a method to automatically recover errors in a brain-computer interface (BCI). As with all non-perfect information transfer mechanisms BCIs suffer from errors. However, error recovery and correction in BCI has largely been neglected.

BCI research is in its infancy with continual exploration of the gamut of functional brain signals in search for the few that are reliable and robust enough for communication [24, 9, 12, 79, 59, 41, 22, 83, 74]. A majority of the research also focuses on the technical side of signal classification in the noisy brain. Signal processing researchers are turning their tools towards the EEG domain as more data sets are available to them. Many algorithms have been looked at to increase the classification rate of both evoked potential-based BCIs and spontaneous EEG BCIs. These methods include simple linear methods such as averaging EEG signals over multiple trials of the same stimulus [25], and correlations between individual trials and predetermined subject averages [6]. Machine learning algorithms have been used to improve BCI accuracy with techniques such as Independent Component Analysis (ICA) for the reduction of artifacts [34] and Support Vector Machines (SVMs) for increasing overall accuracy [42].

BCIs are dependent upon cognitive psychology for functional brain theories to determine the user's intent. In 2000, Falkenstein et al. [43] noted cognitive psychology ignores performance errors possibly because of a lack of error theories and inconsistent error rates across subjects. This neglect has resulted in a minimal set of error theories available for BCI error correction. From this research, the error-related negativity (ERN) wave has been studied and applied the most [43, 47, 78, 15].

Errors in BCI are reduced through intelligent choice organization designed to minimize the occurrence of errors, manual recovery where the user activity selects a choice to recover from an error (e.g. undo), and automatic recovery by a computer (e.g. identifying brain signals associated with errors or applying domain knowledge to the choice such as with a spell checker). If a BCI only presents two choices to a user (e.g. yes/no or on/off) the error recovery

mechanism can directly correct selection to the opposite choice, while a BCI with 3 or more choices (e.g. a speller) can only verify a correct response or, in the presence of an error, undo the selection and represent the choices. In this chapter the state-of-the-art error recovery methods are reviewed.

3.1 Choice Organization

The organization of choices presented to the user in a BCI can greatly impact the number of errors that occur. Techniques to reduce the errors based on the interface include: reducing the number of choices presented to the user, ordering the choices based on their frequency in the domain (e.g. letter frequency), and organizing the choice structure based on the probability of correct selection and rejection to reduce the impact of incorrect selection and incorrect rejection.

In creating a BCI user interface, care must be made in the presentation of choices. For example, in the chess domain a poor binary interface presents choices for all 64 board positions to select a piece and again to place a piece in order to make a move. In this case, the user must accurately select 2 choices and reject 10 choices to make a move — 6 choices to select 1 of 64 positions on the 8×8 board, ($log_2(64) = 6$), 1 choice is selected and 5 are rejected, in order to make a move two positions on the board are selected, the first determines the piece and the second determines the square to place the selected piece, therefore 2 correct selections and 10 correct rejections are required. A better interface only presents valid moves.

For example, only considering valid opening moves reduces the number of choices to 4 ($log_2(10) + log_2(2) = 4$) — in the opening there are ten valid pieces, eight pawns and two knights, and there are only two valid positions to place a selected pawn or knight, see Figure 3.1 for example openings. Reducing the number of choices in this way reduces the number of wrong choices selected, and may additionally reduce the time to select the correct choice if it is incorrectly rejected because less choices need to be cycled through. This type of choice organization also affects other BCIs, such as spellers and icon selection.

Perelmouter and Birbaumer [52] described a method for using a modified binary Huffman tree in their Thought Translation Device (TTD) scanning interface. A binary tree created using the Huffman algorithm optimizes the representation of finite messages in message ensembles based on the probability each message will occur in an ensemble [31]. For example, if messages are letters and ensembles are texts in the English language, then the Huffman algorithm will place the most frequently used letters near the top of the tree so that less branches need to be traversed to select them.

In Perelmouter and Birbaumer's modified Huffman algorithm, the tree is built using letter frequency in English, in addition to the probability of a single correct selection (p) and the probability of a single correct rejection (q) to optimize a scanning interface. A scanning interface presents the user with a single choice that can be actively selected, e.g. making the SCP's amplitude greater then a threshold, or passively rejected, e.g. by not doing anything for a specified

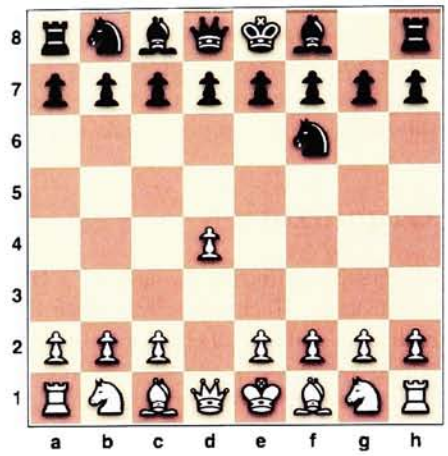


Figure 3.1: Example chess opening moves. Screen capture from Vektor3 3.1.6.

time out period.

In the modified binary Huffman tree algorithm they impose the criteria that the "left child's frequency is not greater than the right child's frequency," and when two trees are combined the frequency of the new tree is the sum of the probabilities weighted by p for the left child and q for the right child:

$$F_{new} = (p \times F_{leftChild}) + (q \times F_{rightChild}) \tag{3.1}$$

By building the tree in this way all left children are selections and right children are the new choices to present because the left child was rejected.

In addition to arranging the alphabet, Perelmouter and Birbaumer also present a method of adding a weighted delete function to the Huffman tree based on the user's error rate, see Section 3.2.1 for details. The modified binary Huffman tree algorithm was presented with a rigorous proof of its optimality, however, no experimental evidence was presented.

The choices in a BCI should be organized to reduce errors based on the domain and the user's performance; however, choice organization will rarely eliminate all errors. The remainder of this chapter reviews error recovery mechanisms employed in current BCIs.

3.2 Manual Error Recovery

Manual recovery involves the user actively selecting an undo choice or actively confirming the selection. Manual recovery is robust across users and only requires the user to be able to select a choice (undo or yes/no for response verification (RV)). Manual recovery's disadvantage is it reduces the system's overall bit rate with the extra choice or confirmation. In addition, the process of selecting the undo may cause more errors.

3.2.1 Undo

When a user selects the undo choice the BCI application reverts to the state before the last selection, e.g. an undo in a BCI speller is similar to the backspace key in a text editor. Typically, undo only reverts non-undo commands; e.g. if the letters *p*, *i*, and *n* have been selected, after two undos the BCI will revert to the state with only *p*. This functionality allows the user to revert the BCI state iteratively back to its initial state. Otherwise, if undo reverted undos the user could not go back more than one state.

The undo is the simplest error recovery mechanism because it can be presented to the user in the same manner as the main choices. The requirement of an extra choice for the user to reject is undo's main disadvantage. This extra choice reduces the speed of the BCI because presenting it takes time and the user may erroneously select it incorrectly requiring the undo to be undone by manually redoing the original action. Placing the undo in the interface follows the same considerations discussed in Section 3.1: Choice Organization.

Perelmouter and Birbaumer [52] present a method for the optimal placement of a backspace (undo) by weighting it according to the user's error rate in a binary Huffman tree. In this way, the backspace will be available to the user only as frequently as the user requires it, reducing the impact on the BCI's speed and the probability of erroneously selecting the undo. Perelmouter and Birbaumer showed when the probabilities of selection and rejection are less than or equal to 0.952, the backspace key is the most frequent 'character' and is placed at the tree's top-level.

3.2.2 Manual Response Verification

Manual response verification is employed when the communication accuracy is more important than the bit rate [82]. In manual response verification the user's selections are explicitly verified by an opposite target trial. When the user selects a choice the result of that choice is the response of the BCI (e.g. selecting the character 'b' causes the BCI to respond with a 'b' on the screen). To verify the response, the BCI presents the user with a yes/no trial. If the user selects yes then the response is correct, if no is selected the response is an error. The yes/no choices are positioned in opposite to the selection to reduce potential bias from the position (potentially by the signal processing) in the choice structure. For example, in a binary interface if the user selects the left target on the first trial, on the response verification trial the left target is no and the right target is yes.

Wolpaw et al. [82] analyzed manual response verification with offline data from one-dimensional and two-dimensional cursor controlled BCIs. In the one-dimensional task the cursor started from the screen center and moved towards the top or bottom of the screen depending on the amplitude of the 8–12 Hz mu rhythm or 20–25 Hz beta rhythm recorded over the sensorimotor cortex. In the two-dimensional task the cursor started from the center of the screen and moved in both directions (like a mouse cursor). 2-D vertical movement was

controlled in the same way as the 1-D task. In the 2-D task, the cursor moved horizontally by the amplitude in a "3 or 5 Hz wide frequency band centered at 10 or 15 Hz from occipital electrodes PO7 and PO8 [82]." Figure 3.2 shows the response verification control flow in the 2-D task [82].

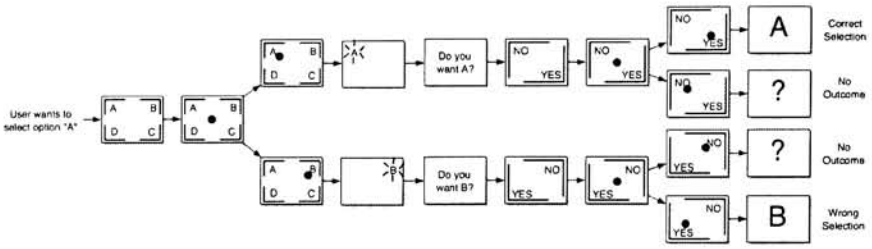


Figure 3.2: Example control flow from Wolpaw et al.'s manual response verification in a 2-D cursor control BCI [82]. The user wants to select A. If the user moves the cursor to A in the upper left corner, the A is selected and the user is then presented with the response to verify. In this verification task No appears in the position of the letter selected to reduce positional bias.

The potential accuracy resulting from using response verification in a 1-D task, C_{1D}, is the probability that two single trials are correct divided by the sum of the probability two single trials are correct and the probability two single trials are incorrect [82]. This is shown in Equation 3.2, where p is the single-trial accuracy, i.e. p is the probability a single selection is correct.

$$C_{1D} = \frac{p^2}{p^2 + (1-p)^2} \tag{3.2}$$

From Equation 3.2, the increase in accuracy from using response verification is greater with more accurate single trial selections (p). Wolpaw gives examples of potential increases when "p is 0.55, or 55%, C_{1D} is expected to be only 60%; if p is 75%, C_{1D} is expected to be 90%; and if p is 91%,C_{1D} is expected to be 99%."

The 2-D task accuracy C_{2D} using response verification is similarly derived in [82] and is dependent on the single trial accuracy of hitting one of the four corner targets p_q and probability of hitting the yes or no in the verification phase p_h.

$$C_{2D} = \frac{p_q p_h}{p_q p_h + (1-p_q)(1-p_h)} \tag{3.3}$$

Wolpaw notes Equation 3.2 is a special form of Equation 3.3 where the "first and second trials have the same accuracy (e.g. when both have the same two-target ... format)." The increase in accuracy from using response verification is impressive when accuracy to select the yes/no, p_h, is much greater than the accuracy to select the four targets p_q. For example, when p_h is 90% and p_q is 50% the new accuracy is 90% [82].

As mentioned by Wolpaw et al. [82], manual response verification does provide increases in accuracy, however, the overall speed of the system is decreased by the addition of the verification trial and the occurrence of undecided trials. According to Wolpaw, "the RV procedure would only become worthwhile in terms of information transfer rate if the second, verification trial were much faster than the first trial." A fast verification trial is one of the goals of automatic response verification presented in the following section.

3.3 Automatic Error Recovery

In automatic error recovery the computer detects errors and performs an undo or selects the correct choice without confirming with the user. Automatic recovery does not reduce the bit rate as manual error recovery does because it does not present the user with an extra choice; however, because auto recovery systems are not perfect, too many errors in correction can frustrate the user and bring the overall bit rate down.

A goal of automatic error recovery in BCI is to utilize the users' brain signals to determine if an error has occurred. In this vein, two signals have been investigated to automatically correct errors: error-related negativity (ERN), which occurs when the user realizes a mistake was made, and the error potential, which occurs when the user is presented with a response that signifies an error. These signals are reviewed below.

3.3.1 Error-Related Negativity (ERN)

The error-related negativity (ERN) is an event-related brain potential that occurs when a subject makes a mistake in a performance task. The ERN consists of a negative deflection, located fronto-centrally, which occurs 80 ms after the subject incorrectly responds to a stimulus [47]. The ERN is also referred to as the Ne (error negativity) in literature. The ERN is sometimes followed by a positive wave Pe that peaks 300 ms after stimulus and may indicate conscious processing and updating to correct and prevent future errors [47].

Blankertz et al. [15, 14] and Parra et al. [51, 50] have experimented with recognizing the ERN in button pressing experiments where subjects were instructed to press a button if an event occurred. Both researchers report successfully reducing errors between 15% and 20%. However no researcher has experimented with the ERN in a task that does not require physical movement. Therefore, it is uncertain how well the ERN will extend to BCI error correction.

3.3.2 Error Potential

Schalk et al. [71] described a positive deflection that occurred 180 ms after incorrect selection over Cz. The potential occurred in offline analysis of 1–D cursor control to select either 'Yes' or 'No' After the word was selected the screen blanked for 80 ms and the word selected was presented to the user. The

four users achieved accuracies of 93.9, 93.1, 85.5, and 85.4% in the original test. Using the error potential for automatic error correction the users' accuracies would have been 95.3, 93.2, 88.4, and 90.6% respectively. In addition, using automatic correction would have increased the BCI's bit rate by 3, 0, 4, and 21 bits/min for the four users.

It is unclear why and how this error potential occurs as it is not mentioned in cognitive psychology literature. In addition, it seems contradictory to the established ERN ERP, which occurs over Cz at 80ms after incorrect selection and is negative, however, as noted above, ERN studies have not been performed without physical movement while the error potential occurred in data from a BCI task. More experiments are required to validate the error potential.

3.4 Conclusion

Manual error recovery mechanisms are useful when high accuracy is required, however, they can slow the BCI's information transfer rate because of the extra choice the user is required to use or respond to. Automatic error recovery mechanisms are ideal because they do not slow the overall transfer rate of the BCI and can increase the accuracy. However, automatic error recovery techniques are in their infancy with ERN having robust characteristics and a large body of cognitive psychology research behind it, but no evidence for its utility in non-physical performance tasks. The error potential was shown in an offline study to increase the accuracy without decreasing the bit rate, however, its robustness remains to be proven.

The rest of this thesis focuses on using the P300 signal, a well established brain signal introduced in Section 2.4.2, for automatic error recovery.

Chapter 4

Automatic Error Recovery Using P3 Response Verification

A method to improve BCI accuracy by automatically recovering errors based on brain signals that occur when the computer provides feedback to a user's selection is described. Experimental results from an evoked potential BCI used to control items in a virtual apartment are presented, and the existence of the evoked potential P3 component in responses to controlled goal items is shown. A reduced response exists when items are accidentally controlled. The presence of the P3 signal in responses to goal items means that it can be used for automatic error correction. Offline experiments were run and with a theoretical mean improvement in recognition from 78% to 85%, there was a statistically significant improvement ($P < 0.008$, Wilcoxon signed rank test) in accuracy of 3% using a correlation algorithm.

4.1 Introduction

Recent research has shown brain-computer interfaces to be an alternative control method for those with severe motor impairment caused by diseases such as Amyotrophic Lateral Sclerosis (ALS) as well as by spinal cord injuries [83, 48, 52, 12, 35]. Many algorithms have been looked at to increase the classification rate of both evoked potential-based BCIs and spontaneous EEG BCIs. These methods include simple linear methods such as averaging EEG signals over multiple trials of the same stimulus [25] as well as correlations between individual trials and predetermined subject averages [6]. Machine learning algorithms have been used to improve BCI accuracy with techniques such as Independent Component Analysis (ICA) for the reduction of artifacts [34] and Support Vector Machines (SVMs) for increasing overall accuracy [42].

This page intentionally not left blank.

There is another way to increase the accuracy of a BCI system. Extra information embedded in the signal can be used to make decisions. This possibility has not been heavily studied, since increasing accuracy by using more information may actually decrease the speed of a system and the end bit rate.

Wolpaw et al. suggested the use of response verification (RV) for applications that need a high accuracy more than they do a high bit rate [82]. Response verification occurs when the classification of a choice for an individual has been made and the individual is then asked to verify that choice by making a further choice. Two types of errors are possible in a BCI: missing the correct choice (a false negative error) or making a choice accidentally (a false positive error). RV may improve accuracy by reducing the occurrence of false positive mistakes in the data. Experimentally, response verification has led to higher accuracies at the cost of a lower overall bit rate [82].

This thesis proposes a method to use automatic error recovery to increase accuracy without decreasing the bit rate of a BCI system. Other methods have been suggested to automatically correct errors. Schalk et al. [71] performed automatic response correction using an error potential they identified in offline data analysis that occurred over Cz after incorrect selection in a 1-D cursor control task. An average 2.4±2.2% increase in accuracy was achieved with response correction. However, Schalk states the error potential requires more experimentation to validate its robustness as it is unclear why and how it occurs and is unidentified in cognitive psychology literature.

Parra et al. perform automatic response error recovery for users accidentally hitting a mouse button in a visual discrimination task [50]. They use the error-related negativity (ERN) event-related potential for automatic recovery. This potential is known to occur when subjects mistakenly respond to a stimulus and realize that they've made a mistake. Unfortunately, the ERN signal is difficult to use in a BCI since subjects do not normally "accidentally" make a mistake — instead a combination of faulty signal classification and faulty subject control over EEG signals combine to cause mistakes. Since the ERN event-related potential is difficult to use and the error potential is unproven, this thesis proposes an error recovery method that depends on the use of the evoked potential P3 component. The P3 component of the evoked potential was independently reported by Chapman and Bragdon [16] and by Sutton et al. [75] and is well established in cognitive psychology. This component is a positive wave peaking at around 300–400 ms after task-relevant stimuli. While the P3 is evoked by many types of paradigms, the most common factors that influence it are the frequency of stimulus occurrence (less frequent stimuli produce a larger response) and task relevance.

After offline analysis of BCI data collected for controlling items in a virtual apartment, it was discovered that the P3 signal exists after a goal item has been controlled and a reduced response occurs after a control mistake. The difference between the P3 signal for goal item responses and mistakes may be used for automatic error recovery and it will increase the bit rate of the system as long as the majority of corrections correct mistakes rather than cancel goal control.

Offline experimental results indicate that with a theoretical possible increase in accuracy from 78% to 85%, automatic response verification can increase the accuracy to maximum of 80%. This improvement is shown using two simple and well-understood algorithms: peak picking and correlation. The simplicity of the algorithms and resulting improvements show the robustness of the effect. It is likely that more complicated machine learning algorithms may improve classification results. Since automatic error recovery is based on the response to controlling of an item and not on the algorithm used in classification for BCI control, the possibility exists for using this technique in multiple types of BCIs.

4.2 The VR apartment experiment

The virtual reality (VR) apartment experiment compared the similarities and differences between the evoked potential P3 component in VR and looking at a computer monitor while using an online P3 signal-based BCI. Results show that the P3 signal is robust between different environments, although the system bit rate was fairly low at 13.49 bits/minute [5]. Offline analysis shows that an evoked potential P3 component exists in responses to goal selections while a reduced signal exists for false positive mistakes.

4.2.1 Experimental Setup

The experimental setups for the VR and non-VR environments were almost identical except that the VR environment is displayed in a head-mounted display (HMD). Five objects or commands could be controlled by the user in the virtual apartment: a lamp, a stereo system, a television set, a Hi command, and a Bye command, see Figure 4.1. The lamp, stereo, and television all worked as toggle switches to turn the items on/off. The Hi and Bye commands made a three-dimensional graphic figure appear (for Hi) or disappear (for Bye). All responses to commands were visual — for instance musical notes appeared over the stereo when the stereo was on.

A sphere associated with each controllable object blinked in the environment; when visible, it had a semi-transparent red coloring. Semi-transparency was used so that blinking spheres would be less distracting to subjects concentrating on one specific sphere for a task. Approximately once per second, a stimulus was provided when the sphere on a randomly chosen item appeared. The stimulus would last for approximately 250 ms. The P3 response occurs for task-relevant stimuli. To make the red sphere flashes on the controllable object task-relevant, subjects had to count the flashes on a goal item.

Seven electrode sites were arranged on the heads of nine subjects with a linked mastoid reference. Sites Fz, Cz, Pz, P3, P4, as well as an upper and lower vertical electrooculographic (EOG) channels were used from the International 10-20 system of placement [32]. For online recognition and analysis, EOG artifacts were regressed out of the signals of interest using the algorithm by Semlitsch [72]. The EEG signals were amplified using Grass amplifiers with

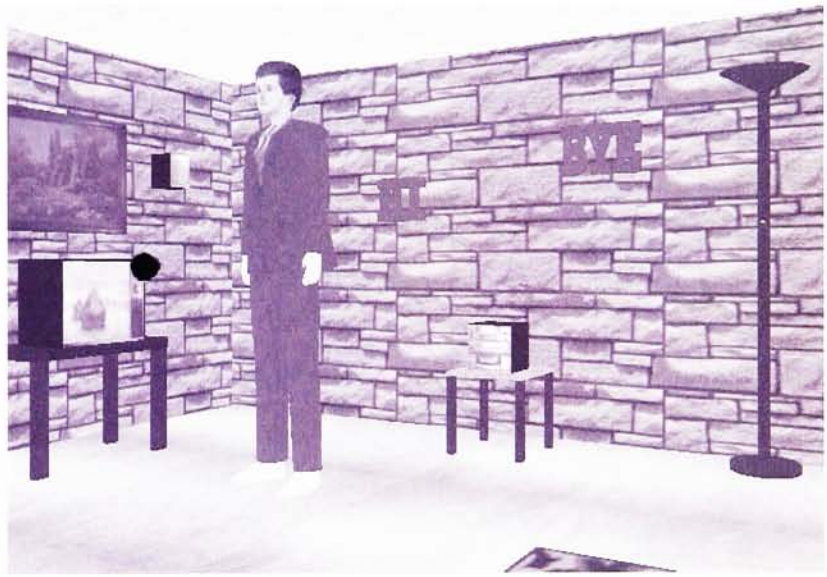


Figure 4.1: A sample scene from the virtual apartment. The television, stereo, HI sign, BYE sign, and lamp are all controllable items. In this scene, a red sphere on the television set is blinking.

an analog bandwidth from 0.1 to 100 Hz. Electrode impedances were between 2 and 10,000 Ω for all subjects. The data were recorded continuously and saved to a file.

4.2.2 The Experiment

The experiment consisted of four tasks:

1. **Calibration**: The subject counted the number of sphere flashes located on the virtual lamp on a monitor.
2. **VR condition**: The subject was fully immersed in the virtual apartment while wearing a HMD.
3. **Monitor condition**: The subject looked at the virtual apartment on a monitor.
4. **Fixed Display condition**: The subject looked at the virtual apartment on a fixed screen inside of the HMD.

The Calibration task was used to train a signal processing algorithm on a particular subject's P3 signal response. A total of 300 stimulus presentations were presented to each subject. In this task, subjects were told to count only the lamp sphere flashes; thus, in this task only the lamp sphere flashes were task-relevant and these flashes should have caused a P3 response. Since the spheres flashed randomly over the five controllable items, 60 ± 5 lamp flashes occurred over the course of five minutes.

Tasks 2–4 were accomplished in a randomized block order and lasted for approximately 5 minutes (250 stimulus presentations with the sphere flashed randomly on items). These tasks involved online single trial classification of the P3 signal in order to control the different items in the apartment. The time taken for these trials depended on how many items were controlled, as subjects received feedback for each item for which the signal classification algorithm classified the trial as a P3 trial.

Due to the difficulty of signal classification, false positive mistakes (accidental control) were possible as well as true goal control. In tasks 2–4, the subjects received English instructions at the bottom of the screen indicating what goal to achieve and each subject attempted to achieve that goal by counting the number of flashes on the sphere located on that particular goal item. During each task, the goal was chosen randomly and the subject tried to achieve the goal for up to 50 presentations of the goal stimulus. When the goal was achieved, an action involving visual feedback occurred in the virtual apartment (for example, the room was lightened when the light was turned on). During the waiting period for this visual feedback (1.5 seconds), no new stimuli were presented. Then, the next goal was randomly chosen.

While the experiment involved online classification and feedback, an offline analysis was done to compare the obtained P3 signals between different conditions. Only epochs with a maximum vertical EOG signal of less than 50 μV were used. This reduced the possibility of EOG contamination of the averages. From this analysis it was discovered that the P3 signals obtained in the different environments were not significantly different from each other. Figure 4.2a shows this result.

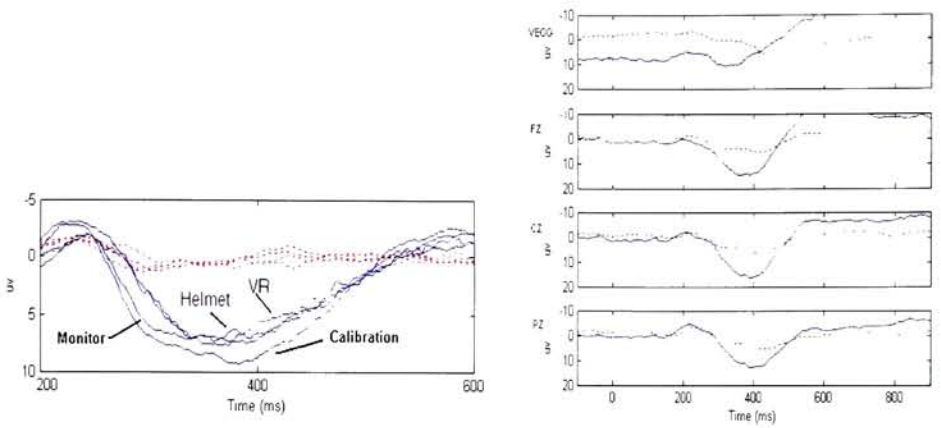


Figure 4.2: a) The grand averages for control of goals at site Pz (solid lines) shown with the grand averages for non-goals (dashed lines) in each experimental condition. b) The grand averages over nine subjects for responses to goal item control (solid lines) and mistakes in item control (dashed lines).

4.2.3 The Appearance of the Evoked Potential P3 Component

Since all items have a discrete control, evoked responses where items are successfully controlled may be examined. The P3 signal occurs when subjects choose a goal item and a reduced signal occurs when a mistake in classification is made. Grand averages showing this result appear in Figure 4.2b. Since subjects could have blinked or moved directly after an item was controlled, all trials used in the grand averages had a maximum recorded vertical eye movement of less than 50 μV. The grand average for vertical eye movement is shown and while it is not flat, there are no peaks around 400 ms when the maximal P3 signal appears.

The maximum signal for Fz and Cz are slightly larger than the signal for Pz. It is possible that subjects found controlling items in a virtual apartment to be "novel" and that would lead to a more frontal response. Controlled goal items are task relevant because the subject achieves control and may go on to the next task. It is hypothesized that false positive mistakes do not generally cause this response since subjects do not always have to correct errors, so they ignore them.

4.3 Methods

4.3.1 Using an Evoked-Potential for Error Recovery

The method to use an evoked potential for response verification (RV) is similar to the response verification procedure described in Wolpaw et al. [82]. However, instead of prompting the user to verify the selection with a Yes/No trial, the system automatically keeps the selected item only if the evoked potential is present.

If p is the single-trial accuracy for selecting goals and rejecting non-goals, and q is the accuracy for goals kept when selected non-goals are discarded during response verification, then the predicted RV proportion decided, D, is:

$$D = pq + (1 - p)(1 - q) \tag{4.1}$$

This is the sum of the probability of keeping a selected goal (correct trial) and keeping a selected non-goal (incorrect trial). The predicted RV accuracy, C, is the probability a decision is correct and may be represented as:

$$C = \frac{pq}{pq + (1 - p)(1 - q)} \tag{4.2}$$

When prompted in manual RV, the user's q probability is similar to the the user's single-trial accuracy, but in automatic response verification the user is not actively responding. This makes it difficult to predict the q probability. In the experiment, the q was found to be less than the user's single-trial accuracy, with a mean of 61% ($P < 0.02$).

4.3.2 Offline Analysis

The epochs corresponding to the nine subjects' previously recorded responses from the virtual apartment experiment Monitor condition were labeled as goal and non-goal responses. For example, if a goal response was the lamp lighting when it was the goal, a non-goal response would be the TV turning on when the lamp was the goal. These were further divided into responses with P3 signals present and P3 signals absent, yielding four categories of responses: Goal P3 Present, Goal P3 Absent, Non-goal P3 Present, and Non-goal P3 Absent.

Two algorithms - peak picking and correlation were considered separately to recognize the evoked potential P3 component in the Fz, Cz, and Pz channels. Vertical-electrooculographic (VEOG) artifacts were regressed out of the signals of interest using the algorithm by Semlitsch [72]. Response epochs containing signal amplitude greater than 90 μV were ignored. This allowed for noisy response trials to be ignored. Ignoring a response resulted in items remaining triggered. For example, if the lamp was lit, it would stay lit.

If the P3 signal was present on Fz, Cz, Pz, or any of the three channels, then the response would be kept. This technique was used for both algorithms compared.

Peak Picking

Peak picking is a simple algorithm to recognize a P3 signal using the difference between the minimum and maximum amplitude in a trial. A trial with a prototypical evoked potential P3 component contains a large peak around 300–400 ms and peak picking recognizes a P3 signal when the amplitude difference is greater than or equal to a specified voltage difference between the minimum and maximum voltage points within a specified time window as expressed by:

$$\max(x) - \min(x) \geq \tau \begin{cases} 1 & : \text{ P3 Present} \\ 0 & : \text{ P3 Absent} \end{cases} \tag{4.3}$$

where x is a vector that represents the data for a single evoked potential (EP) response and τ represents the threshold voltage difference required to accept a P3 signal. For recognition, the time window with the best results was between three and six hundred milliseconds. The voltage difference threshold was optimized for each subject using data from the Monitor condition.

Correlation

Correlation maybe looked at as template matching when correlation is performed between single trials and templates of a P3 and non-P3 signal. EP responses were correlated with P3 and non-P3 signal averages created from the Calibration task using the following formula:

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{4.4}$$

where x is a vector that represents the data for an EP response, y is a vector that represents the P3 or non-P3 signal average to compare against, $cov(x,y)$ is the covariance of x and y, and σ is the standard deviation of the appropriate signal.

To determine if an EP response contains a P3 signal, the correlation between the EP response trial and the P3 signal average was compared with the EP response trial and non-P3 signal average according to the algorithm:

$$\begin{array}{lll} \text{if} & \rho_{P3} > \tau \text{ and } \rho_{P3} \geq \rho_{nonP3} & \text{then P3 Present} \\ \text{else} & \text{P3 Absent} & \end{array} \tag{4.5}$$

ρ_{P3} is the correlation between the EP response and P3 signal average, ρ_{nonP3} is the the correlation between the EP response and non-P3 signal average, and τ is the threshold set to determine the desired amount of correlation. The threshold was varied in experiments to yield the best results.

4.4 Results

Figure 4.3 illustrates the response verification (RV) accuracy versus single-trial accuracy for correlation. The figure shows that keeping a response if the P3

was found increased the overall accuracy for each subject using the correlation algorithm. Peak picking behaved similarly.

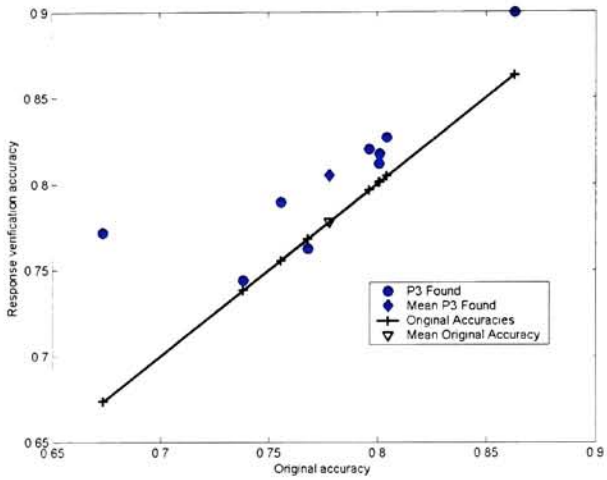


Figure 4.3: Response verification accuracy vs original accuracy using the correlation method. Solid line shows the original accuracy vs itself. Items above the line increased in performance.

Table 4.1 compares the best RV accuracies achieved by the two algorithms against the original accuracy. An overall RV accuracy of 80±5% ($P < 0.05$) was achieved by peak picking and 81±4% ($P < 0.008$) by correlation. All P values were derived using the Wilcoxon signed rank test.

In addition, Table 4.1 presents the theoretical best accuracy that can be achieved in this experiment when correcting mistakes. Overall, it was possible to improve the accuracy from 78% to 85% if all false positives were correctly rejected without introduction of new false negatives. This indicates that 15% of the errors were caused by missed goals, while 7% were caused by mistakes, and of those mistakes, 3% were corrected using automatic response verification in this experiment.

In order to calculate the bit rate, the bits/trial equation used was [82]:

$$B_t = \log_2 N + P \log_2 P + (1 - P) \log_2 \frac{1 - P}{N - 1} \tag{4.6}$$

where there are N possible selections, each of which are equally probable, and an accuracy probability of P. The bit rate B_m, is the bits/trial B_t, multiplied by the average number of trials/minute T_m:

$$B_m = B_t T_m \tag{4.7}$$

In the virtual apartment experiment N was 5 and T_m was 12 trials/minute.

Table 4.1: Change in accuracy from the original accuracy when keeping responses with P3 signals using the peak picking and correlation algorithms to recognize P3 signals. The best theoretical accuracy occurs if all selected goals are kept while all selected non-goals are rejected (P calculated with the Wilcoxon sign rank test).

Subject	Original	New Accuracy		Theoretical
	Accuracy	Peak Pick	Corr.	Best Accuracy
1	80	78	81	87
2	80	82	82	91
3	86	89	90	91
4	80	81	82	88
5	67	77	77	80
6	80	83	83	85
7	76	77	79	84
8	77	77	76	82
9	74	74	74	76
Mean	78	80	81	85
Std	5	5	4	5
		$P < 0.04$	$P < 0.008$	

Table 4.2 shows the bit rate for each subject. The overall increase in bit rate was 1.26 bits/minute and the greatest increase was 4.01 bits/minute by subject five. Subject five's change reflects the benefit of using automatic response error recovery. Subject five's original accuracy was 67% with a large amount of errors from mistakes, 13%. Through response error recovery the subject achieved 77% accuracy. See Appendix B for further results.

4.5 Discussion and Future Work

The existence of the evoked potential P3 component for goal responses has been shown to aid in increasing the accuracy of a BCI system without decreasing the bit rate. While a gain in accuracy of 3% appears small, it is statistically significant and on par with related automatic error recovery methods. In addition, the original experiment was intentionally tuned to reduce selected non-goals at the expense of rejected goals, which resulted in only 7% of the error occurring from selecting non-goals. Therefore, the best increase in accuracy possible using any response verification method with this data set is 7%.

The results appear robust and should be extended to a wider variety of algorithms to show possible system improvements. In addition, the technique needs to be used in an online experiment in order to show how error recovery affects later subject responses. Since individuals adapt to BCI systems, offline results often differ from online results.

Automatic error recovery for a BCI is useful and it is possible that the recovery may be used in multiple types of BCI systems. While automatic error

Table 4.2: Change in the bit rate from the original bit rate when keeping responses with P3 signals using the peak picking and correlation algorithms to recognize P3s. The best theoretical accuracy occurs if all selected goals are kept while all selected non-goals are rejected.

Subject	Original	New Bits/minute		Theoretical
	Bits/minute	Peak Pick	Corr.	Best Bits/minute
1	14.45	13.44	14.98	17.80
2	14.23	15.40	15.40	20.48
3	17.66	19.53	19.81	20.38
4	14.45	14.72	15.27	18.86
5	9.09	12.87	13.10	14.30
6	14.62	16.13	15.74	16.92
7	12.37	13.13	13.91	16.46
8	12.92	13.15	12.69	15.62
9	11.63	11.88	11.88	12.63
Mean	13.49	14.47	14.75	17.05
Std	2.38	2.32	2.22	2.65

recovery works for the discussed experimental paradigm, the existence of the evoked potential P3 component may be tied to the experimental paradigm used.

This page intentionally not left blank.

Chapter 5

Chess BCI Platform for Experimentation

An experimental platform for error recovery was created to address the limitations of the VR apartment offline data analysis study. The VR apartment error recovery study was constrained in three ways. First, the initial online experiment was tuned to minimize false positives, thus limiting the number of false positives for automatic error recovery. Second, subjects were able to ignore non-goal responses, therefore subjects may have received more stimuli from true positives than false positives. Finally, subjects' reactions to incorrect automatic response verification were absent from offline analysis, whereas incorrect recovery online will cause a change in the interface that the subjects may react to.

The chess domain was selected for the platform environment because chess is engaging and requires the user's full attention, chess itself is a slow game, therefore the user may not feel as frustrated with the system compared to a BCI speller, and chess requires 100% accuracy. The selection of the chess domain was also motivated by a collaboration with researchers at the Institute of Medical Psychology and Behavioral Neurobiology (IMPBN) at the University of Tübingen in Tübingen, Germany.

IMPBN developed a brain-computer interface (BCI) to enable completely paralyzed, or locked-in, patients the ability to communicate using slow cortical potentials. Currently this system is used for communication between patients' care-givers and researchers, however, one patient has requested an interface to play chess, which he currently plays with his nurses using eyebrow movements.

The RIT AI Lab did not have the equipment necessary to record EEG at the collaboration's inception, while IMPBN had a full BCI system, but lacked the technical resources necessary to develop a chess BCI. Through the collaboration, IMPBN's system would be used for online experimentation and their patient would have access to a chess BCI.

The following sections describe the chess BCI platform for experimentation

and the collaboration with IMPBN.

5.1 Patient's Brain-Computer Interface

The patient, MS at IMPBN, uses eye-brow movements and the Thought Translation Device, an EEG Brain-Computer Interface (BCI) developed by Birbaumer et al. in BCI2000 [69], to communicate. The Thought Translation Device (TTD) uses patient controlled slow cortical potential shifts for binary selections. Slow cortical potentials (SCP) are slow low frequency changes detected with scalp-recorded EEG that occur over 0.5-10.0 seconds [83].

In operation, the TTD presents a user with two letter banks representing left and right alphabet halves. The user has a time limit to shift his SCP past a positive or negative threshold to select between the left and right letter banks, respectively. In Birbaumer's experiments the time limit was between 4.5 and 6.0 seconds.

When a letter bank is selected, it is split in half and presented to the user in the same fashion as above until the user selects a final letter. If the user does not select a bank after two successive tries a 'go back' function appears as an option to go back to the previous letter bank. After training, patients with 65-90% spelling accuracy can write 0.15-3.0 letters/min [9, 83].

5.2 IMPBN's Chess BCI Requirements

IMPBN had the following requirements, all of which were fulfilled:

1. Play chess against the user at various levels of expertise.
2. Allow the user to play against a human using the same computer.
3. Allow the user to play against a human across a network, preferably interfacing with an online chess program such as www.chessclub.com
4. The program must interface with the user's current brain-computer interface (BCI2000).
5. The chess application must run under Windows and interface with Borland C++ Builder 5.5 code.

5.3 Chess BCI Design

Figure 5.1 depicts a high level chess brain-computer interface design. The chess application is separated into a Graphical User Interface (GUI) that displays the chess board and any information related to the game. The computer chess engine is the program that plays chess against the patient. This design is based on the majority of chess application designs that are in turn derived from XBoard and GNU/Chess.

GNU/Chess is a text based command line chess program where a player issues move commands to the engine through standard input and receives opposing moves through standard output. XBoard is a graphical application that interfaces to GNU/Chess through pipes. XBoard provides a graphical chess board where moves are made using a mouse.

GNU/Chess was ported to Windows and with it came the Winboard graphical chess interface, which communicates using the text command protocol. Because Winboard is an open interface and Windows is ubiquitous, many chess engine programmers, more interested in strategy and tactics than GUIs, wrote their engines to use the Winboard protocol. There are now roughly 120 chess engines that support the Winboard protocol [76].

As with many de facto standards, there are other competing protocols. The most successful competitor is Universal Chess Interface (UCI), created for the commercial chess engine Shredder. However, there are chess interfaces that support both Winboard and UCI.

Most graphical chess interfaces, such as Winboard and Arena, allow two engines to play against each other. This is beneficial as it allows engine authors an easy way to test their engines against other engines. The chess BCI was developed to leverage this ability. As shown in Figure 5.1, the proxy chess engine interacts with the chess GUI as a normal chess engine, but proxies the patient's chess moves, communicated via the BCI, to the GUI.

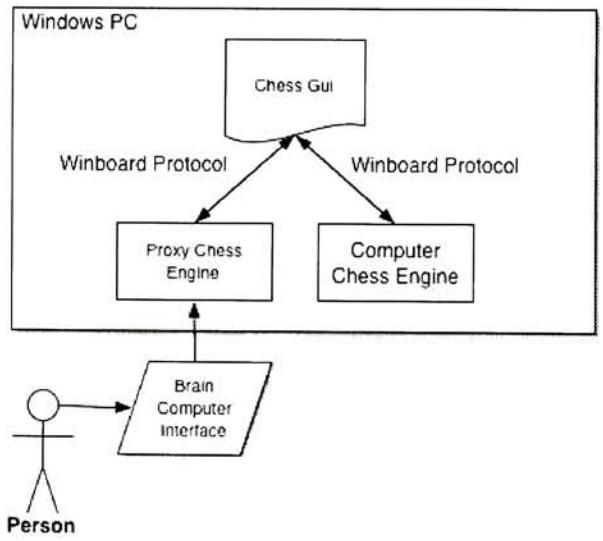


Figure 5.1: Chess Brain-Computer Interface High Level Design

5.4 Implementation

The chess BCI platform consists of a Proxy Engine to communicate the moves selected by the BCI to the chess program, modifications to the Winboard source code for BCI, and modules implemented for BCI2000. Winboard was used because it already had the functionality to play user against engine, engine against engine, and either online with a transparent interface. Combining Winboard with the Proxy Engine facilitated the fulfillment of IMPBN's requirements.

5.4.1 Proxy Engine

The Proxy Engine is a socket based client-server application written in C. The client and server transmit text chess moves from standard input and output bidirectionally to each other. When using Winboard for the chess interface, Winboard starts the server as if it was a generic chess engine. The BCI application then starts the client. When the BCI, client, and server are ready the server tells Winboard the "chess engine" is ready using the Winboard protocol.

By separating the Proxy Engine into a client-server model, the chess interface and opponent chess engine can run on a separate machine from the BCI. This is beneficial because the BCI and chess engine are processor intensive, and the BCI may have its own interface that competes for display resources with the chess interface on a single monitor. In addition, by decoupling the BCI from the chess interface using the proxy engine in this way, a different BCI client can easily be connected to the chess interface with a minimum amount of integration.

5.4.2 Winboard Modifications

The source code to Winboard was modified to present move stimuli for a P300 BCI. The pieces and squares flashed different colors to evoke the P300 and create a move. Valid pieces and squares to flash are determined through hooks created into the Winboard game model and move validation code. The code allows piece and square stimuli to be grouped to create different stimuli presentation patterns. For example, each square can be flashed individually, or squares can be grouped into rows, which the user can drill down through to select the individual piece position.

In the modified Winboard interface, the stimuli (pieces and squares) can flash in three different modes: constant, cycle, or random. In constant mode, each stimuli is flashed with the same color. In cycle mode, a fixed cycle of colors is created, e.g. RED, YELLOW, GREEN, and repeated modulo the stimuli. For example, if there are ten stimuli, S1 ... S10, the flashing would be S1-RED, S2-YELLOW, S3-GREEN, S4-RED, S5-YELLOW, ... S10-YELLOW. In random mode, each stimuli flashed with a random color. These three modes were implemented to test novelty, predictability, and habituation on the P300 response from the stimuli.

Finally, Winboard was modified to operate using a flexible stage flow to control when events occurred in the interface. The stage system controls when

the system pauses for the user to think, flashes the stimuli for piece then square selection, moving the the piece to the selected square, and transmitting the move to the opponent. Stages can be easily inserted and removed without affecting other stages in the system.

5.4.3 BCI2000 Modules Implemented

Chess GUI Interface

A chess GUI was implemented as a BCI2000 application module to provide the BCI fine grained control of the GUI not easily available in Winboard.

BCI2000 [69] is a BCI application written by the Wadsworth Institute in Borland C++ and provides a flexible framework with separation of signal acquisition, signal processing, and the user interface. As mentioned previously, the patient's BCI was implemented in BCI2000.

An oddball chess application module was created to calibrate the system with the user's template P300 and non-P300 signals. In the oddball application a random piece appeared around the chess board in random locations. The pieces shown could be varied, e.g. king, queen, knight, and rook. A chess game GUI was also created including piece movement animation. The open source "Simple Chess Engine" created by Rémi Coulom [17] was adapted to BCI2000 to provide a chess game model and valid move generation for stimulus display.

EEG Source Acquisition Module

The RIT AI lab purchased a Tucker-Davis Technologies (TDT) RA16 biosignal amplifier for EEG acquisition after the start of the IMPBN collaboration. An EEG source acquisition module was written to use the RA16 with BCI2000. The TDT ActiveX driver was ported from Visual C++ to Borland C++ for the BCI2000 acquisition module as BCI2000 is coded in Borland C++. The acquisition module was implemented using a multi-threaded producer-consumer model. In the implementation, data was stored in a list and placed on a queue. A list pool was created to reduce memory creation and usage.

5.5 Discussion and Future Work

After the first milestone was released to IMPBN they stopped responding, possibly because the student we were communicating with graduated — by this time the AI lab purchased an EEG amplifier, which the chess application was ported to using BCI2000. After the experimental platform was created it was found the amplifier transfered data too slowly for real-time multi-channel BCI, therefore online experiments were not performed. The slow transfer speed was identified after numerous tests of the system and hardware, and was found to be caused by the Tucker-Davis Technologies driver reinitializing the USB interface after the call to transfer data. At the time of this writing Tucker-Davis

Technologies developed a work around for this deficiency, and the AI Lab will use this chess BCI platform for future experiments.

Chapter 6

Conclusion

This thesis presented a method to automatically recover errors by detecting the P3 component for response verification, where the presence of the P3 component verifies the selection was correct. An overview of BCI was given including a review of alternative error recovery strategies. The method proposed was validated in an offline experiment with data from a virtual apartment controlled by a P3 BCI. In the experiment, the P3 error recovery method improved the accuracy and bandwidth of the tested BCI using a simple P3 detection algorithm. This provides evidence for the utility of the P3 error recovery method and justification for further exploring its application in BCI.

6.1 Future work

The P3 error recovery method was verified offline in an experiment where the original false positives were purposefully limited at the expense of false negatives, thus reducing the impact of automatic response verification. If more false positives were available in the data set the method could have more significantly increased the accuracy and bandwidth of the system. With this, and the generality that false positives and false negatives are diametrically opposed in minimization, an experiment to test the effect of minimizing false negatives in the signal processing while using automatic response verification to minimize false positives should be performed. In addition, it needs to be determined if too many automatic response verifications slow the system due to signal processing or incorrectly classifying the response.

Because offline experimental results generally differ from online experimental results the P3 error recovery method should be tested in an online experiment.

Because the P3 is evoked by any task relevant stimulus this error recovery method should also work in other BCIs and should be tested with them to prove its robustness across BCI applications.

Schalk et al. [70] noted the error potential they witnessed may be a P3 signal, which somewhat contradicts the findings of this study. In Schalk's work,

the error potential was found when the user was given feedback on an incorrect selection, while in this study the P3 was found when the user was given feedback on a correct selection. Schalk hypothesized the error potential may be an oddball reaction to the errors because the subjects generally had very low error rates. It may be that as users become more competent with the BCI system P3s move from occurring on successes to occurring on errors. Further research is required for clarification.

Finally, the P3 error recovery method proposed should be tested in a multimodal error recovery system where it works in conjunction with other error recovery methods. For example, using manual response verification when automatic response verification cannot decide if a P3 has occurred may increase the overall accuracy with minimal or no reduction in transmission speed. In addition, detecting the ERN when the P3 does not exist, and detecting the P3 when the ERN does not exist, can reduce incorrect error classifications by increasing the system's classification confidence.

Appendix A

Accuracy Equations

This appendix presents additional accuracy equations used in analysis of automatic response verification from the point of view of using selection and rejections of items instead of probability. The intention is to provide additional information to aid in the understanding of automatic response verification.

In the following, *selected* and *rejected* refer to the user selecting and rejecting stimuli, while *kept* and *discarded* refer to the computer keeping and discarding selected goals during response verification. Selected and kept and rejected and discarded are synonyms respectively, but this terminology was used to clarify which stimuli and goals are being referred to.

Accuracy

A generic equation for accuracy in a BCI is:

$$Accuracy = \frac{SelectedGoals + RejectedNonGoals}{TotalStimuli}$$

$$where:$$

$$TotalStimuli = SelectedGoals + SelectedNonGoals + RejectedGoals + RejectedNonGoals$$

Where SelectedGoals are the number of selected goal items, SelectedNonGoals are the number of selected non-goal items, etc.

Improving Accuracy with Response Verification

Response verification (RV) will only improve accuracy by discarding user selected non-goals. RejectedGoals are not changed because when a goal item is rejected the user has missed the item and no feedback is given by the computer. In addition, RV can erroneously reject previously selected goals and therefore reduce the accuracy.

Selections are preserved in RV as follows:

$$SelectedGoals = KeptGoals + DiscardedGoals$$

$$SelectedNonGoals = KeptNonGoals + DiscardedNonGoals$$

$DiscardedGoals$ and $KeptNonGoals$ need to equal 0 to achieve the theoretical best accuracy RV can produce in a system.

Accuracy of the RV trial

The response verification is itself a trial with an accuracy. The accuracy of the RV trial is also the q probability referred to in Section 4.3.1, and is defined as:

$$\text{Accuracy of RV} = \frac{KeptGoals + DiscardedNonGoals}{SelectedGoals + SelectedNonGoals}$$

New Overall Accuracy with RV

The new accuracy of a system using response verification is:

$$\text{New Accuracy with RV} = \frac{KeptGoals + DiscardedNonGoals + RejectedNonGoals}{TotalStimuli}$$

Categorization and Accuracy

Selections need to be categorized in some way for response verification to work. Manual response verification requires the user to classify the selections into *yes* or *no* categories. In automatic response verification with the P3, ERN, or error potential the selections are categorized as signal *present* or signal *absent* The following holds in this case:

$$\begin{aligned} SelectedGoals &= SelectedGoalsSignalPresent + SelectedGoalsSignalAbsent \\ SelectedNonGoals &= SelectedNonGoalsSignalPresent + SelectedNonGoalsSignalAbse \end{aligned}$$

Depending on what the signal indicates (success or error) one of two methods is used to determine which category of selections to keep, and which to discard:

I: If the signal indicates success, then keep only selections with the signal *present*:

$$accuracy = \frac{KeptGoalsSignalPresent + DiscardedNonGoalsSignalAbsent + RejectedNonGoals}{TotalStimuli}$$

II: If the signal indicates error, then keep only selections with the signal *absent*:

$$accuracy = \frac{KeptGoalsSignalAbsent + DiscardedNonGoalsSignalPresent + RejectedNonGoals}{TotalStimuli}$$

Case I applies to RV with the P3; case II applies the RV with the ERN and the error potential.

Indeterminate Categorization

It is possible the signal categorization algorithm can respond with present, absent, and indeterminate. For instance, the correlation algorithm proposed in Section 4.3.2 can be modified as follows to produce this indecisive ternary output:

$$\begin{array}{llll} \text{if} & \rho_{P3} > \tau \text{ and } \rho_{P3} \geq \rho_{nonP3} & \text{then} & \text{P3 Present} \\ \text{elseif} & \rho_{nonP3} > \tau \text{ and } \rho_{nonP3} \geq \rho_{P3} & \text{then} & \text{P3 Absent} \\ \text{else} & \text{Indeterminate} & & \end{array} \tag{A.1}$$

ρ_{P3} is the correlation between the response signal and P3 template, ρ_{nonP3} is the the correlation between the response signal and non-P3 template, and τ is the threshold set to determine the desired amount of correlation. In this way, there is greater confidence a P3 is either present or absent. In addition, the ternary automatic RV system can be paired with a manual RV system, deferring to a potentially more accurate but slower method. See Section B.3 for offline data analysis on the VR apartment experiment using indeterminate categorization.

Accuracy with RV and Ignored Responses

Finally, it is often the case that some responses cannot be used for response verification. For example, in the offline analysis in Section 4.3.2 responses with signals greater than 90 μV were ignored. By ignoring responses some SelectedNonGoals may remain; and the accuracy changes because some SelectedGoals not counted in KeptGoals. The new overall accuracy equation is:

$$\text{Accuracy with RV and ignored responses} = \frac{KeptGoals + DiscardedNonGoals + RejectedNonGoals + IgnoredGoals}{TotalStimuli}$$

This page intentionally not left blank.

Appendix B

Data used for Response Verification Analysis of the Virtual Apartment Experiment Screen Condition

Data was analyzed offline from nine subjects participating in a BCI virtual apartment experiment to determine the efficacy of response verification using the P300 signal. Overall results are shown in Chapter 4, while results from individual subjects are shown in this appendix. Please see Chapter 4 for a discussion of the experiment and main results.

This appendix consists of three sections reporting best results from the Peak Pick, Correlation, and Correlation Indeterminate algorithms to determine the existence of the P300 signal during response verification. The data was regressed and epochs with signals greater than 90 μVolts were discarded as reported in Section 4.3.1.

The Correlation Indeterminate algorithm is similar to the Correlation algorithm described in Section 4.3.2, however, if the response did not correlate above the threshold with either the P3 signal average or the non-P3 signal average the trial was considered *indeterminate*, and would not be affected by response verification. See Section B.3 for further details.

B.1 Peak Pick

Table B.1: The number of responses used in subject averages. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

Subjects	Goal Responses	NonGoal Response	Total
1	18	12	30
2	15	19	34
3	3	10	13
4	13	15	28
5	4	24	28
6	2	6	8
7	12	15	27
8	12	11	23
9	3	4	7
Total	82	116	198

Table B.2: Stimuli from original experiment broken into selected, rejected, goal, and non-goal. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

Subjects	Selected Goal	Rejected Goal	SelectedNonGoal	RejectedNonGoal	Total
1	19	24	13	130	186
2	16	14	20	117	167
3	34	16	14	155	219
4	15	20	16	130	181
5	5	35	28	125	193
6	4	19	7	103	133
7	12	26	17	121	176
8	13	32	13	136	194
9	4	34	11	123	172
Total	122	220	139	1140	1621

Table B.3: Discarded EOG: The number of stimuli not used in analysis because their signals were greater than the EOG threshold. Thus these were unaffected by response verification, but were still used to determine the new accuracy after response verification. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

Subjects	Selected Goal	Selected NonGoal	RejectedGoal	RejectedNonGoal	Total
1	1	4	1	20	26
2	1	1	1	5	8
3	31	15	4	44	94
4	2	4	1	22	29
5	1	5	4	15	25
6	2	4	1	20	27
7	0	8	2	47	57
8	1	2	2	13	18
9	1	9	7	34	51
Total	40	52	23	220	335

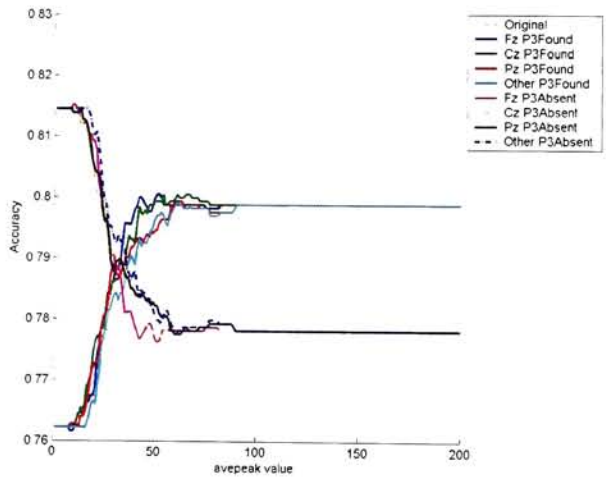


Figure B.1: Accuracies resulting from varying average peak for Peak Pick algorithm. Regressed true, EOG Threshold: 90

Table B.4: Original and response verification accuracies. Verification (q) indicates the accuracy of the response verification trial itself. Theoretical best indicates best accuracy that could have been achieved if all selected goals were kept and all selected non-goals were discarded during response verification. All P values were calculated with the Wilcoxon sign rank test. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

Subject	Original	Keep Response if P3 was found in channel:				Keep Response if P3 was Absent in channel:				Verification (q) if P3 was Found in channel:				Verification (q) if P3 was Absent in channel:				Theoretical Best
		Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	
1	0 80	0 77	0.78	0.77	0.78	0.80	0.79	0 80	0.79	0.44	0.47	0.44	0.47	0.56	0 53	0.56	0.53	0.87
2	0.80	0.82	0 82	0.82	0.82	0.80	0.80	0 80	0 80	0.56	0.56	0.56	0 56	0.44	0.44	0.44	0.44	0.91
3	0 86	0.89	0 89	0.89	0.89	0.86	0.86	0 86	0 86	0 85	0.85	0.85	0 85	0.71	0.71	0.71	0.71	0.91
4	0.80	0.81	0.81	0 81	0.81	0.80	0.80	0 81	0 81	0 55	0.55	0.52	0.52	0.48	0.48	0.52	0 52	0 88
5	0 67	0.77	0.78	0.77	0 77	0.68	0.67	0 68	0.68	0 73	0.76	0.73	0.70	0.18	0.15	0.18	0 21	0 80
6	0.80	0.83	0.83	0.83	0.83	0 80	0.80	0 80	0 80	0 73	0.73	0.73	0.73	0.36	0.36	0.36	0.36	0.85
7	0 76	0.77	0 77	0 77	0.77	0 76	0.76	0 76	0.76	0.52	0 52	0 52	0 52	0.41	0.41	0.41	0.41	0.84
8	0 77	0 76	0 77	0 77	0 77	0.77	0.76	0 76	0.76	0.46	0 50	0.54	0 54	0.50	0 16	0.42	0 42	0.82
9	0 74	0 74	0 74	0 74	0 74	0 74	0.74	0 71	0.74	0 33	0 33	0 33	0.33	0 27	0.27	0.27	0.27	0 76
Mean	0 78	0 80	0 80	0.80	0 80	0 78	0.78	0 78	0.78	0.57	0 58	0.58	0 58	0.44	0.13	0.43	0.43	0.85
Std	0.05	0.05	0.05	0 05	0.05	0.05	0.05	0 05	0 05	0 17	0.16	0.16	0.16	0.16	0.16	0.16	0.15	0 05
Significance P<		0.074	0.055	0 055	0.039	1 000	0.500	0 875	0.875	0.012	0 020	0 012	0 012	0.004	0.004	0.004	0.004	

Table B.5: Original and response verification bits per minute. Verification (q) indicates the bits per minute of the response verification trial itself. Theoretical best indicates the best bit rate that could have been achieved if all selected goals were kept and all selected non-goals were discarded during response verification. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

Subject	Original	Keep Response if P3 was found in channel:				Keep Response if P3 was Absent in channel:				Verification (q) if P3 was Found in channel				Verification (q) if P3 was Absent in channel.				Theoretical Best
		Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	
1	14.45	13.20	13.44	13.20	13.44	14 19	13.94	14 19	13.94	2.50	3.15	2.50	3.15	5.50	4 65	5.50	4 65	17.80
2	14 23	15 40	15.40	15.40	15.40	14 23	14.23	14 23	14.23	5.30	5.30	5.30	5.30	2 64	2.64	2 64	2 64	20.48
3	17.66	19.53	19.53	19 53	19.53	17.66	17.66	17.66	17.66	17 17	17.17	17.17	17 17	10.41	10 41	10 41	10 41	20 38
4	14 45	14.99	14.99	14.72	14 72	14 45	14 45	14 72	14.72	5.11	5 11	4.26	4.26	3.49	3.49	4.26	4 26	18.86
5	9 09	13.10	13.33	13 10	12 87	9.28	9.09	9.28	9.48	11.17	12 46	11.17	9 97	0.02	0 14	0.02	0 01	14.30
6	14 62	16.13	16.13	16 13	16.13	14.62	14.62	14.62	14.62	11.17	11.17	11.17	11.17	1.24	1.24	1 24	1.24	16.92
7	12.37	13.13	13 13	13 13	13.13	12.37	12 37	12 37	12 37	4.29	4 29	4 29	4 29	2.05	2 05	2.05	2.05	16.46
8	12 92	12 69	12 92	13 15	13 15	12 92	12 69	12 46	12 46	2 99	3 86	4 84	4 84	3.86	2 99	2 22	2.22	15.62
9	11 63	11 88	11 88	11 88	11 88	11 63	11 63	11.63	11.63	0 84	0 84	0 84	0 84	0.22	0.22	0 22	0.22	12.63
Mean	13 49	14 45	14 53	14 47	14.47	13 48	13.41	13 46	13.46	6 73	7 04	6 84	6.78	3 27	3 09	3 17	3 08	17.05
Std	2 38	2 36	2 31	2 32	2.32	2 33	2 37	2.36	2.31	5.31	5 33	5.23	5.06	3 21	3 12	3 24	3 17	2.65

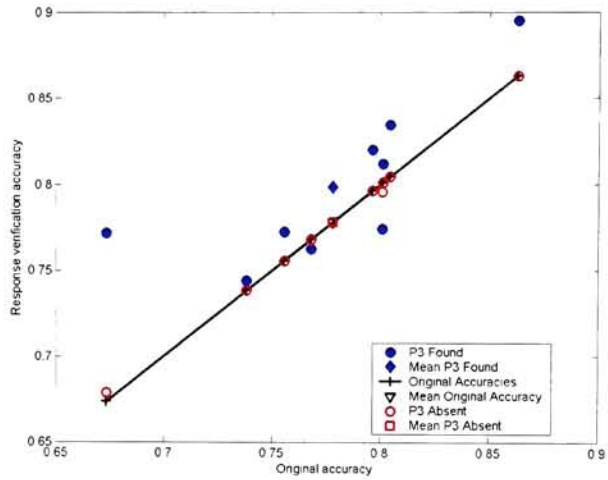


Figure B.2: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Fz). Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61.

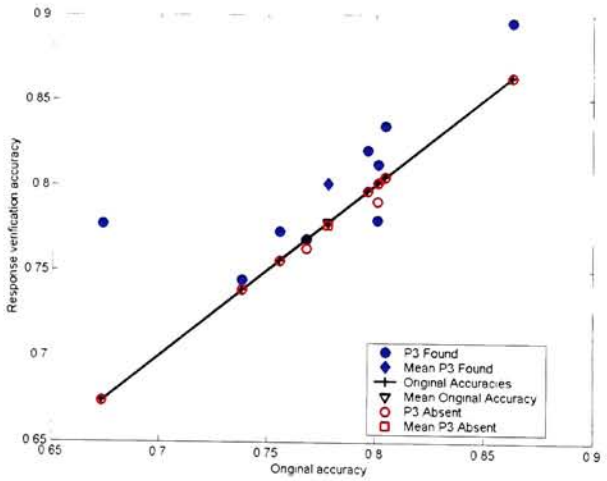


Figure B.3: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Cz). Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61.

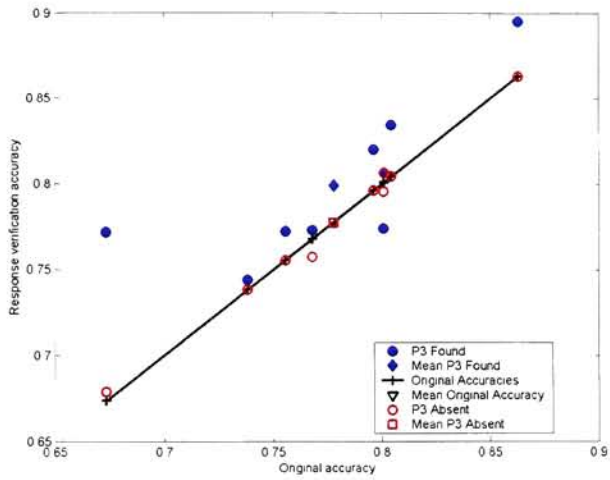


Figure B.4: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Pz). Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61.

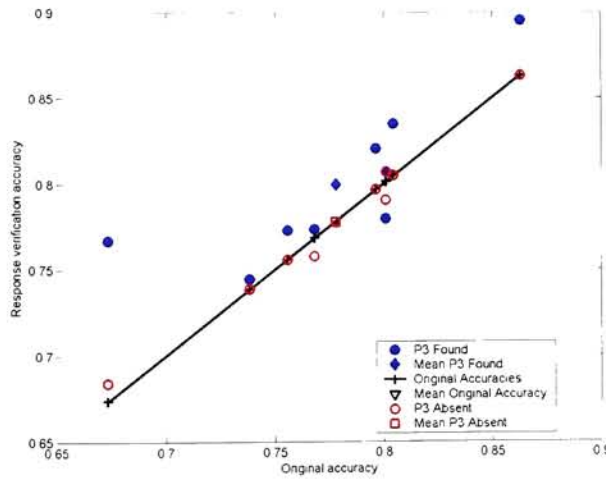


Figure B.5: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Any). Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61.

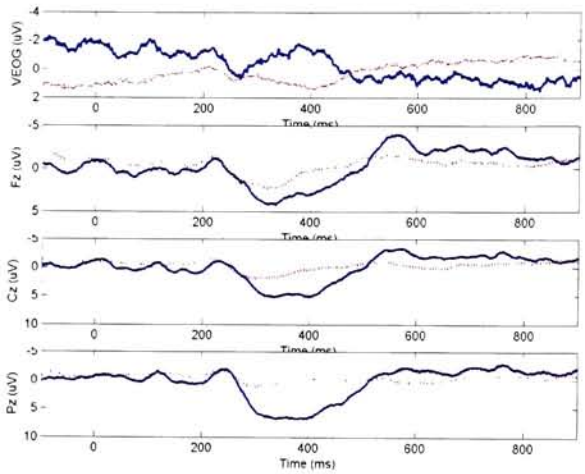


Figure B.6: Across All Subjects Response Verification Trials Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

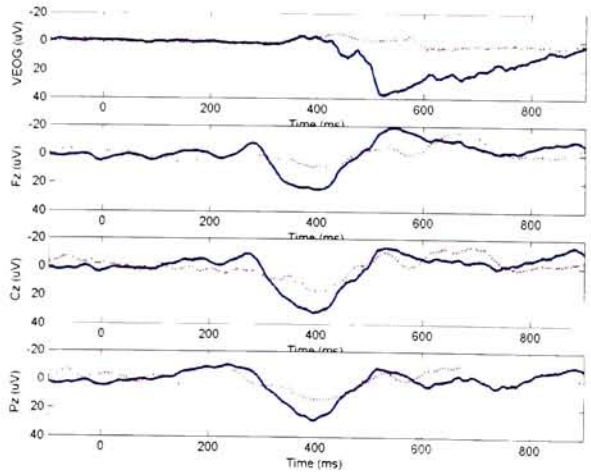


Figure B.7: Response Verification Trials Subject 1 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

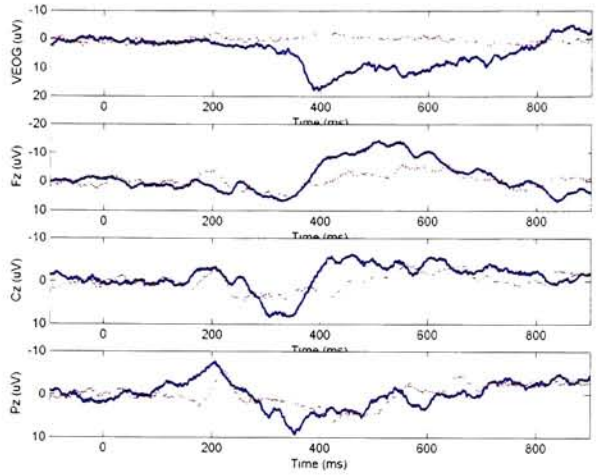


Figure B.8: Response Verification Trials Subject 2 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

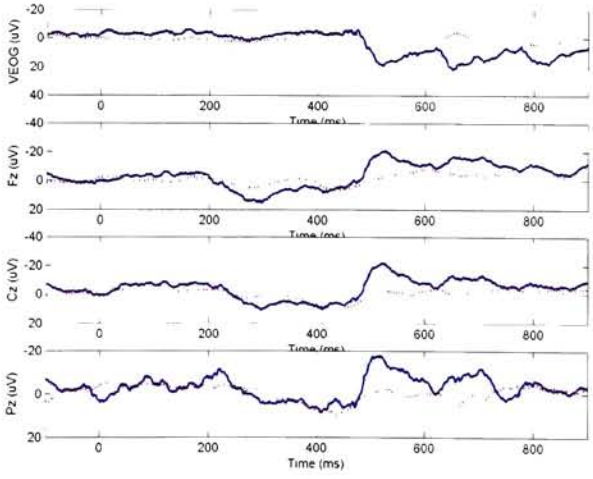


Figure B.9: Response Verification Trials Subject 3 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

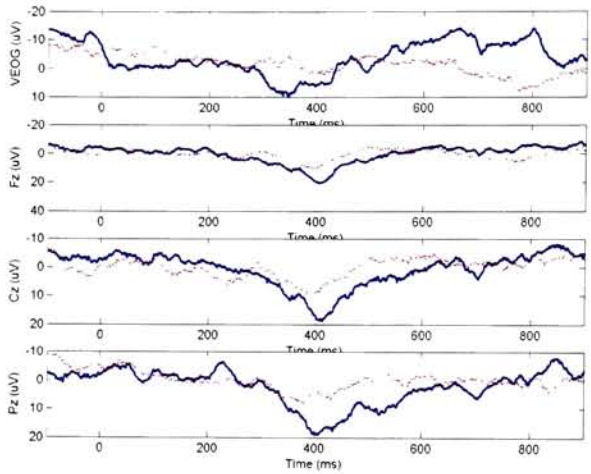


Figure B.10: Response Verification Trials Subject 4 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

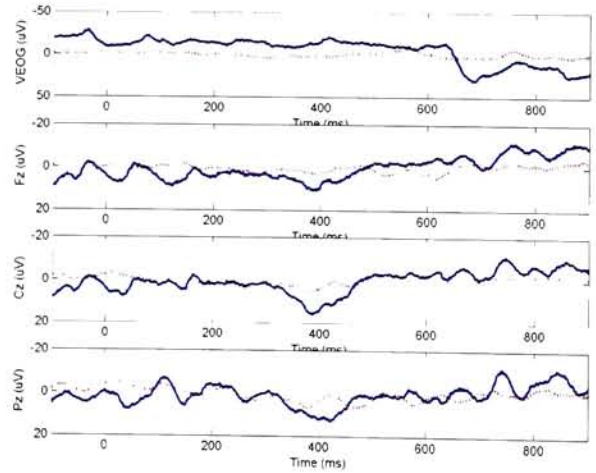


Figure B.11: Response Verification Trials Subject 5 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

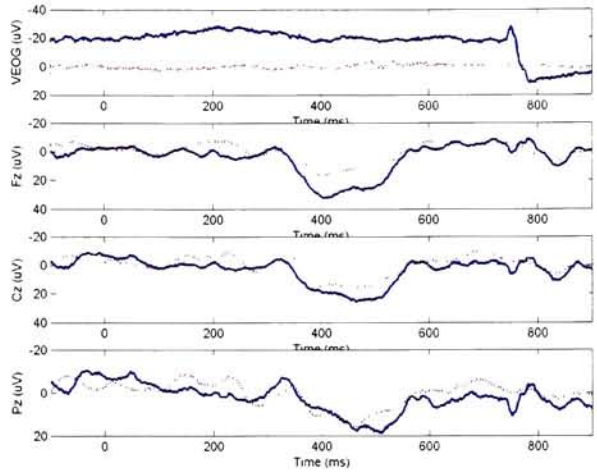


Figure B.12: Response Verification Trials Subject 6 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

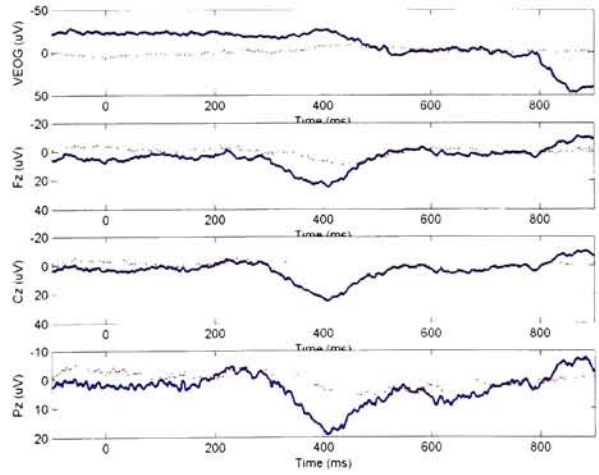


Figure B.13: Response Verification Trials Subject 7 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

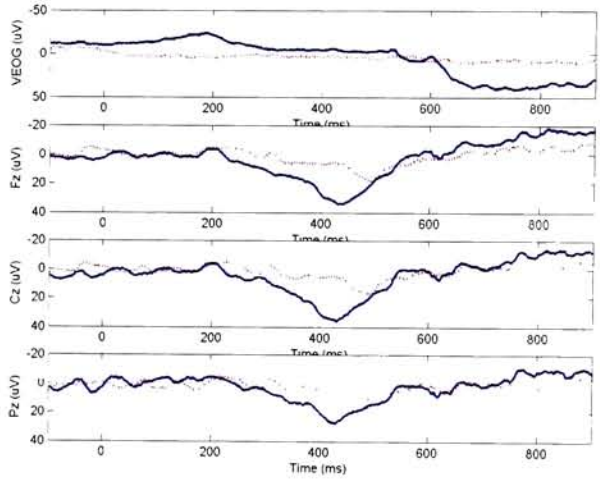


Figure B.14: Response Verification Trials Subject 8 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

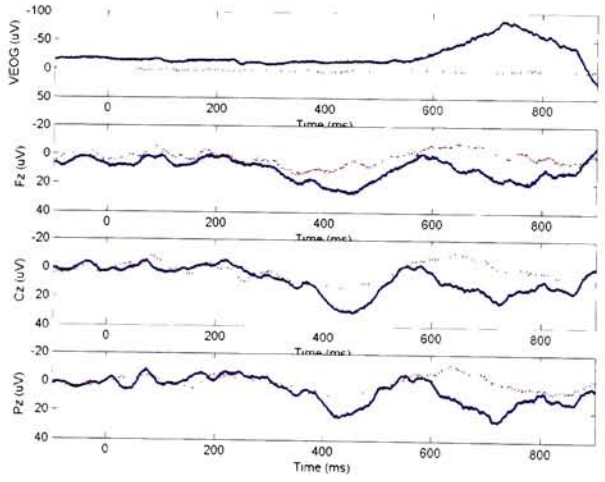


Figure B.15: Response Verification Trials Subject 9 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Peak Pick, Regressed: yes, EOG Threshold: 90, Peak Threshold: 61

B.2 Correlation

Table B.6: The number of responses used in subject averages. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subjects	Goal Responses	NonGoal Response	Total
1	18	12	30
2	15	19	34
3	3	10	13
4	13	15	28
5	4	24	28
6	2	6	8
7	12	15	27
8	12	11	23
9	3	4	7
Total	82	116	198

Table B.7: Stimuli from original experiment broken into selected, rejected, goal, and non-goal. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subjects	Selected Goal	Rejected Goal	SelectedNonGoal	RejectedNonGoal	Total
1	19	24	13	130	186
2	16	14	20	117	167
3	34	16	14	155	219
4	15	20	16	130	181
5	5	35	28	125	193
6	4	19	7	103	133
7	12	26	17	121	176
8	13	32	13	136	194
9	4	34	11	123	172
Total	122	220	139	1140	1621

Table B.8: Discarded EOG: The number of stimuli not used in analysis because their signals were greater than the EOG threshold. Thus these were unaffected by response verification, but were still used to determine the new accuracy after response verification. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subjects	Selected Goal	Selected NonGoal	RejectedGoal	RejectedNonGoal	Total
1	1	4	1	20	26
2	1	1	1	5	8
3	31	15	4	44	94
4	2	4	1	22	29
5	1	5	4	15	25
6	2	4	1	20	27
7	0	8	2	47	57
8	1	2	2	13	18
9	1	9	7	34	51
Total	40	52	23	220	335

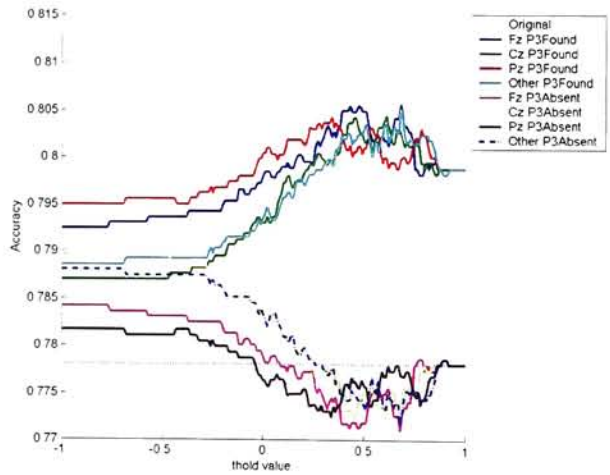


Figure B.16: Accuracies resulting from varying threshold for Correlation algorithm. Regressed true. EOG Threshold: 90

Table B.9: Original and response verification accuracies. Verification (q) indicates the accuracy of the response verification trial itself. Theoretical best indicates best accuracy that could have been achieved if all selected goals were kept and all selected non-goals were discarded during response verification. All P values were calculated with the Wilcoxon sign rank test. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subject	Original	Keep Response if P3 was found in channel:				Keep Response if P3 was Absent in channel:				Verification (q) if P3 was Found in channel:				Verification (q) if P3 was Absent in channel:				Theoretical Best
		Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	
1	0.80	0 79	0.79	0.81	0.81	0.78	0.78	0.76	0 76	0.53	0.53	0.63	0.66	0.47	0.47	0.38	0.34	0.87
2	0 80	0.83	0.82	0.81	0.82	0.79	0.80	0.80	0.80	0.58	0.56	0.53	0.56	0.42	0.44	0.47	0.44	0.91
3	0.86	0.90	0 89	0.89	0.90	0 86	0.86	0.86	0.86	0.88	0.85	0.85	0.88	0.69	0.71	0.71	0.69	0.91
4	0 80	0.82	0.82	0.82	0 82	0.80	0.80	0.80	0 80	0.58	0.58	0.58	0.58	0.45	0.45	0.45	0.45	0.88
5	0.67	0.78	0.78	0.77	0.77	0.67	0.67	0.68	0.68	0 76	0.76	0.73	0.73	0.15	0.15	0.18	0 18	0 80
6	0 80	0.84	0.84	0.81	0.83	0.80	0.80	0.83	0 81	0.82	0.82	0 45	0.64	0.27	0.27	0.64	0 45	0 85
7	0 76	0.79	0.77	0.77	0 79	0.74	0.76	0.76	0.74	0.62	0.52	0 52	0.62	0.31	0.41	0.41	0.31	0.84
8	0 77	0.76	0.76	0.76	0.76	0.77	0.77	0 77	0.77	0 46	0.46	0.46	0.46	0.50	0.50	0.50	0.50	0 82
9	0 71	0.74	0.74	0.74	0 74	0.74	0.74	0 74	0.74	0.33	0.33	0.33	0.33	0.27	0.27	0 27	0.27	0 76
Mean	0 78	0.81	0 80	0.80	0 81	0.77	0.77	0.78	0 77	0.62	0.60	0.56	0.61	0.39	0.41	0.45	0.40	0.85
Std	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.17	0.17	0 16	0.15	0.16	0.16	0 16	0 15	0 05
Significance P<		0 027	0 027	0 008	0 008	0.031	0.250	1 000	0.438	0.055	0.039	0 012	0 020	0.004	0.004	0 004	0 004	

Table B.10: Original and response verification bits per minute. Verification (q) indicates the bits per minute of the response verification trial itself. Theoretical best indicates the best bit rate that could have been achieved if all selected goals were kept and all selected non-goals were discarded during response verification. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subject	Original	Keep Response if P3 was found in channel:				Keep Response if P3 was Absent in channel:				Verification (q) if P3 was Found in channel:				Verification (q) if P3 was Absent in channel:				Theoretical Best
		Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	
1	14 45	13.94	13.94	14 71	14.98	13.44	13.44	12.71	12.48	4.65	4 65	7.41	8 47	3.15	3.15	1.41	0.97	17.80
2	14.23	15.70	15.40	15.10	15.40	13.95	14.23	14.52	14.23	6.10	5.30	4.56	5.30	2.10	2.64	3.22	2.64	20.48
3	17 66	19.81	19.53	19.53	19.81	17.41	17.66	17.66	17 41	18.34	17 17	17.17	18.34	9 61	10.41	10.41	9.61	20.38
4	14 45	15.27	15.27	15.27	15.27	14.19	14.19	14.19	14.19	6.02	6.02	6.02	6.02	2.78	2.78	2.78	2.78	18.86
5	9.09	13 33	13.33	13.10	13.10	9.09	9.09	9.28	9.28	12.46	12.46	11.17	11 17	0.14	0.14	0.02	0.02	14.30
6	14 62	16.52	16.52	14.99	15.74	14.26	14.26	15.74	14 99	15.29	15.29	2.84	7.79	0 26	0 26	7.79	2.84	16.92
7	12 37	13.91	13.13	13 13	13 91	11 64	12 37	12.37	11.64	7 27	4.29	4.29	7 27	0.59	2.05	2.05	0 59	16.46
8	12 92	12.69	12.69	12.69	12 69	12.92	12.92	12.92	12.92	2.99	2 99	2.99	2 99	3.86	3.86	3.86	3.86	15.62
9	11 63	11.88	11.88	11 88	11 88	11 63	11 63	11.63	11 63	0 84	0 84	0 84	0 84	0 22	0.22	0.22	0.22	12 63
Mean	13 49	14 78	14.63	14 49	14 75	13 17	13.31	13 45	13.20	8 22	7 67	6.37	7.58	2.52	2.84	3.53	2.62	17.05
Std	2 38	2 39	2.35	2.25	2 32	2 30	2.32	2 43	2 34	5.86	5 79	5.04	5.05	3.01	3.16	3.47	2.95	2.65

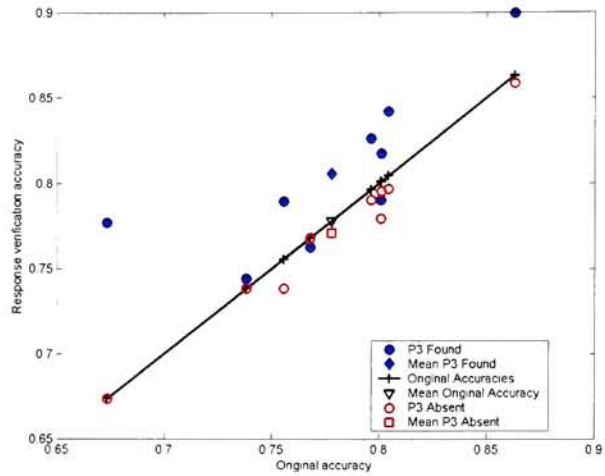


Figure B.17: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Fz). Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

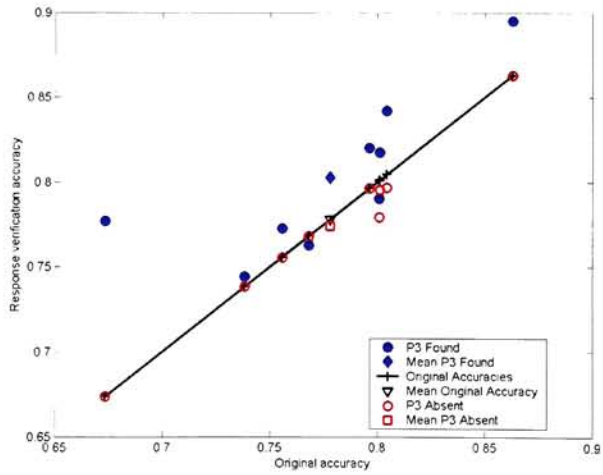


Figure B.18: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Cz). Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

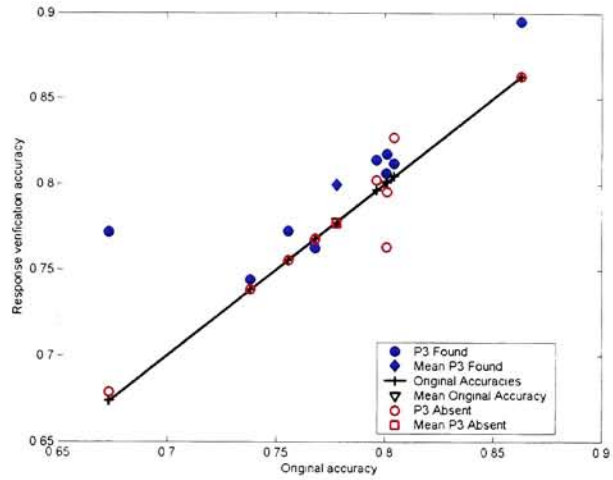


Figure B.19: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Pz). Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

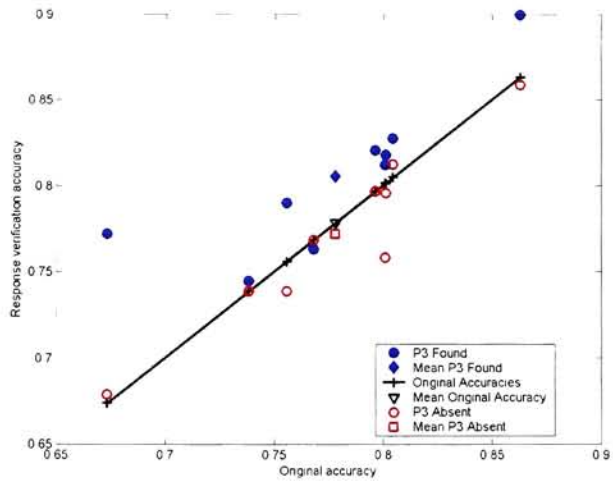


Figure B.20: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Any). Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

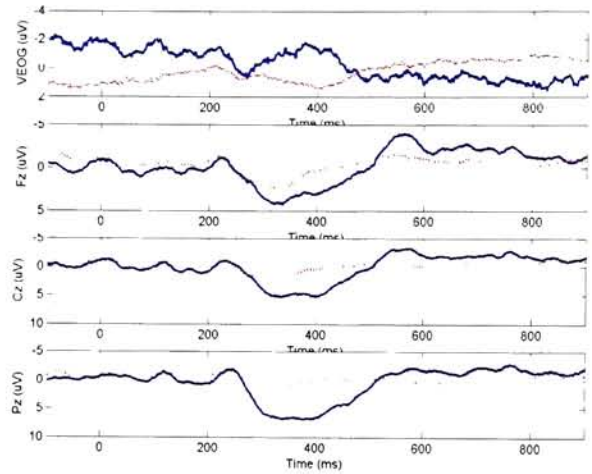


Figure B.21: Across All Subjects Response Verification Trials Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

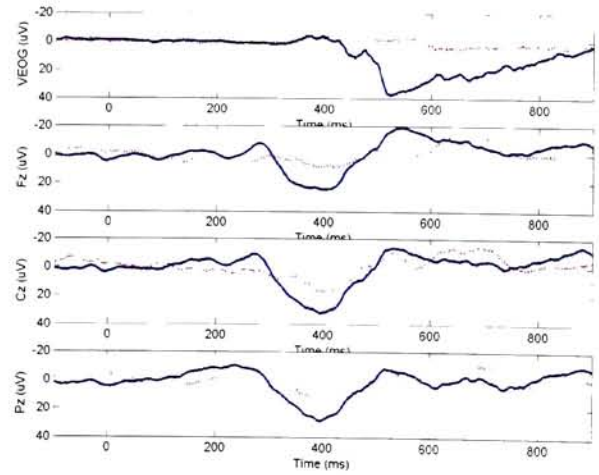


Figure B.22: Response Verification Trials Subject 1 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

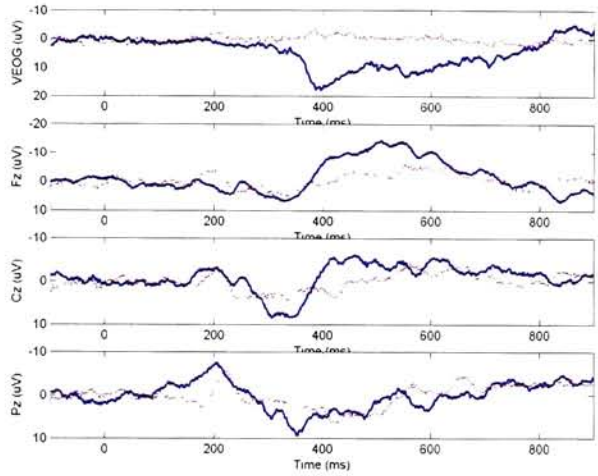


Figure B.23: Response Verification Trials Subject 2 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

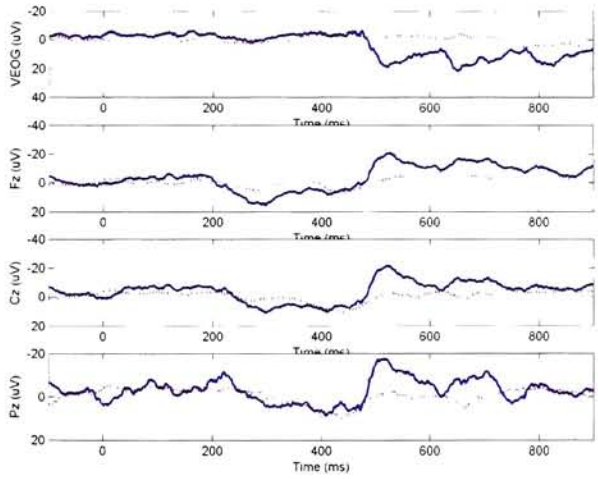


Figure B.24: Response Verification Trials Subject 3 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

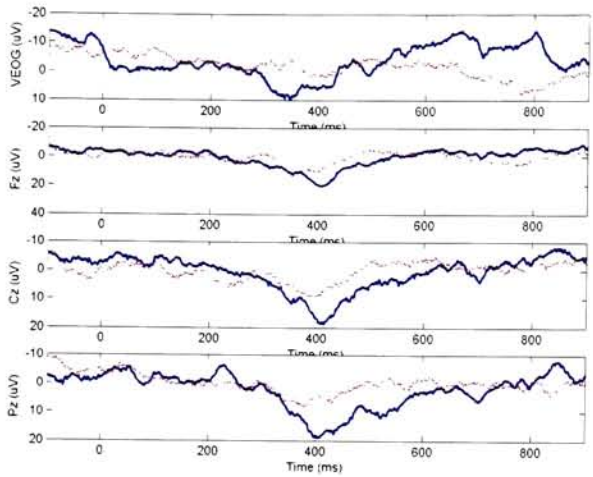


Figure B.25: Response Verification Trials Subject 4 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

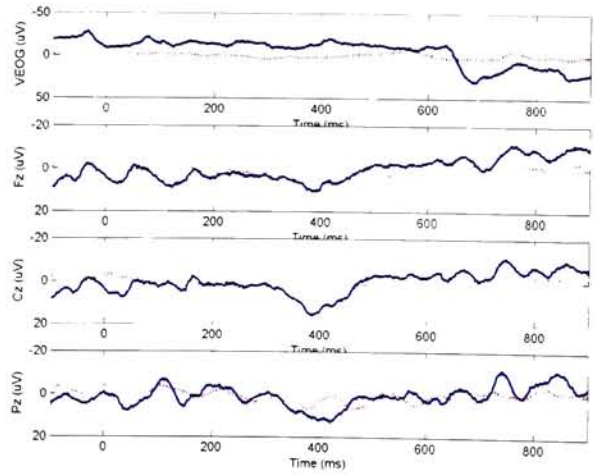


Figure B.26: Response Verification Trials Subject 5 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

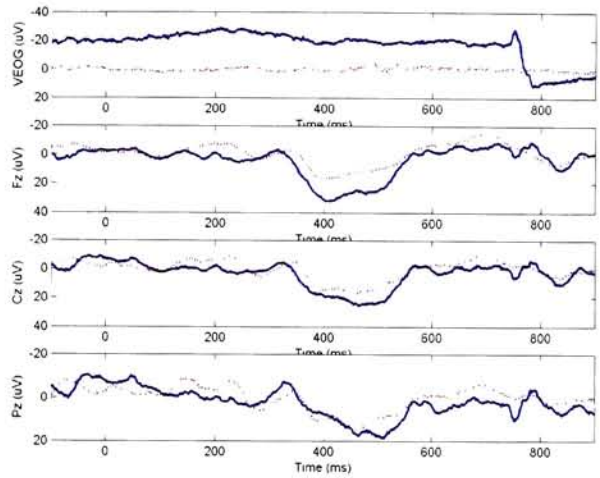


Figure B.27: Response Verification Trials Subject 6 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

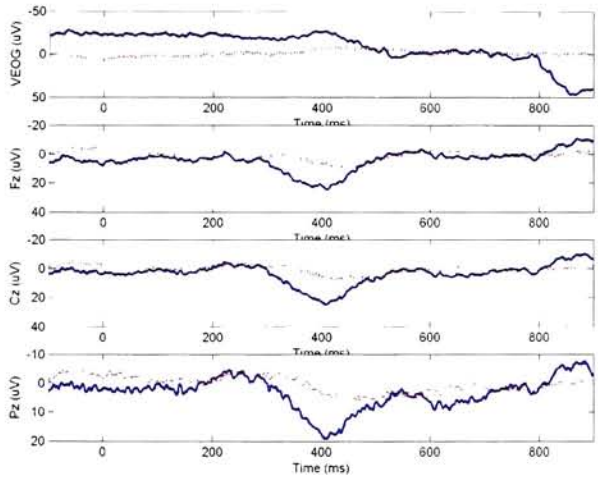


Figure B.28: Response Verification Trials Subject 7 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

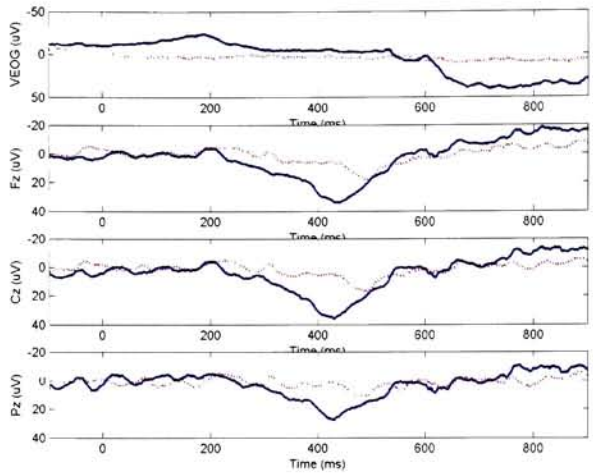


Figure B.29: Response Verification Trials Subject 8 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

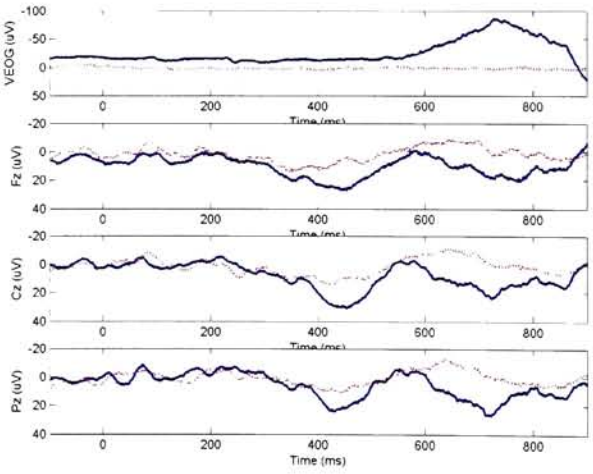


Figure B.30: Response Verification Trials Subject 9 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Fz. Condition: Monitor, Algorithm: Correlation, Regressed: yes. EOG Threshold: 90, Correlation Threshold: 0.68

B.3 Correlation Indeterminate

Correlation may be looked at as template matching where correlation is performed between single trials and a template of a P3 and non-P3 signal. EEG responses were correlated with P3 and non-P3 averages created from the calibration task using the following formula:

$$\rho_{x,y} = \frac{cov(x,y)}{\sigma_x \sigma_y} \tag{B.1}$$

Where x is a vector that represents the data for an EEG response, y is a vector that represents the P3 or non-P3 average to compare against, $cov(x,y)$ is the covariance of x and y, and σ is the standard deviation of the appropriate signal. To determine if an EEG response contains a P3 the correlation between the EEG response epoch and the P3 average was compared with the EEG response epoch and non-P3 average according to the algorithm:

$$\begin{array}{llll} \text{if} & \rho_{P3} > \tau \text{ and } \rho_{P3} \geq \rho_{nonP3} & \text{then} & \text{P3 Present} \\ \text{elseif} & \rho_{nonP3} > \tau \text{ and } \rho_{nonP3} \geq \rho_{P3} & \text{then} & \text{P3 Absent} \\ \text{else} & \text{Indeterminate} & & \end{array} \tag{B.2}$$

ρ_{P3} is the correlation between the EEG response and P3 average, ρ_{nonP3} is the the correlation between the EEG response and non-P3 average, and τ is the threshold set to determine the desired amount of correlation. The threshold was varied in experiments to yield the best results. If the EEG response did not correlate with a P3 or non-P3 average the algorithm is indeterminate. Variable averaging, described in the next section, improved the treatment of indeterminates.

Table B.11: The number of responses used in subject averages. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subjects	Goal Responses	NonGoal Response	Total
1	18	12	30
2	15	19	34
3	3	10	13
4	13	15	28
5	4	24	28
6	2	6	8
7	12	15	27
8	12	11	23
9	3	4	7
Total	82	116	198

Table B.12: Stimuli from original experiment broken into selected, rejected, goal, and non-goal. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subjects	Selected Goal	Rejected Goal	SelectedNonGoal	RejectedNonGoal	Total
1	19	24	13	130	186
2	16	14	20	117	167
3	34	16	14	155	219
4	15	20	16	130	181
5	5	35	28	125	193
6	4	19	7	103	133
7	12	26	17	121	176
8	13	32	13	136	194
9	4	34	11	123	172
Total	122	220	139	1140	1621

Table B.13: Discarded EOG: The number of stimuli not used in analysis because their signals were greater than the EOG threshold. Thus these were unaffected by response verification, but were still used to determine the new accuracy after response verification. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subjects	Selected Goal	Selected NonGoal	RejectedGoal	RejectedNonGoal	Total
1	1	4	1	20	26
2	1	1	1	5	8
3	31	15	4	44	94
4	2	4	1	22	29
5	1	5	4	15	25
6	2	4	1	20	27
7	0	8	2	47	57
8	1	2	2	13	18
9	1	9	7	34	51
Total	40	52	23	220	335

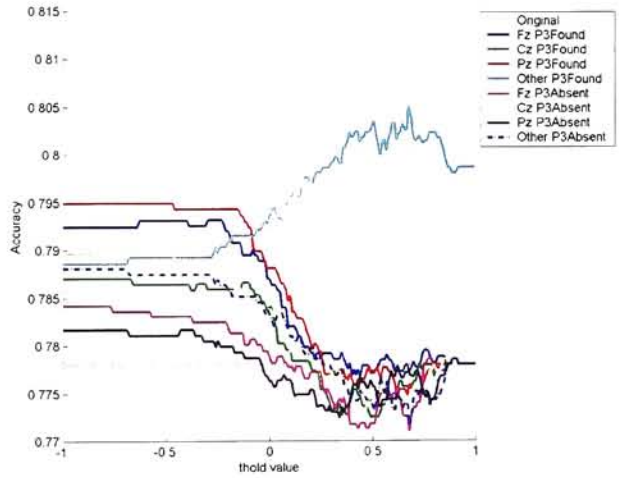


Figure B.31: Accuracies resulting from varying threshold for Correlation Indeterminate algorithm. Regressed true, EOG Threshold: 90

Table B.14: Original and response verification accuracies. Verification (q) indicates the accuracy of the response verification trial itself. Theoretical best indicates best accuracy that could have been achieved if all selected goals were kept and all selected non-goals were discarded during response verification. All P values were calculated with the Wilcoxon sign rank test. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Subject	Original	Keep Response if P3 was found in channel:				Keep Response if P3 was Absent in channel:				Verification (q) if P3 was Found in channel:				Verification (q) if P3 was Absent in channel:				Theoretical Best
		Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	
1	0.80	0.80	0.80	0.81	0.81	0.78	0.78	0.76	0 76	0 59	0.59	0.63	0 66	0 47	0.47	0.38	0.34	0 81
2	0.80	0.80	0.80	0 79	0.82	0 79	0.80	0.80	0.80	0 44	0.47	0.42	0 56	0 42	0.44	0 47	0 44	0 80
3	0.86	0 86	0.86	0 86	0 90	0 86	0.86	0.86	0.86	0.71	0 71	0.71	0 88	0 69	0.71	0.71	0.69	0 86
4	0 80	0.80	0.81	0 80	0 82	0 80	0.80	0.80	0 80	0 48	0 55	0.45	0 58	0 45	0 45	0.45	0.45	0 80
5	0 67	0 67	0.67	0 67	0.77	0.67	0.67	0.68	0 68	0 15	0.15	0.15	0 73	0 15	0.15	0.18	0 18	0 67
6	0 80	0.81	0.80	0.80	0.83	0 80	0.80	0.83	0.81	0.45	0 36	0 36	0 64	0 27	0.27	0.64	0 45	0 82
7	0.76	0.76	0.73	0.74	0.79	0 74	0 76	0 76	0.74	0 45	0 24	0 34	0 62	0 31	0.41	0.41	0.31	0 78
8	0 77	0.75	0.76	0 77	0.76	0 77	0 77	0 77	0.77	0 35	0 12	0 50	0 46	0 50	0 50	0.50	0.50	0 77
9	0 74	0.74	0 74	0.74	0 74	0 74	0 74	0.74	0.74	0.27	0 27	0 27	0 33	0 27	0.27	0.27	0 27	0 71
Mean	0 78	0 78	0.78	0.78	0 81	0.77	0.77	0.78	0 77	0 43	0 42	0 43	0 61	0 39	0 41	0 45	0 40	0 78
Std	0.05	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0 05	0 17	0 18	0 17	0 15	0 16	0 16	0 16	0 15	0 05
Significance P<		1 000	0.875	0.250	0 008	0.031	0.250	1.000	0 438	0.004	0 004	0 001	0 020	0 004	0 004	0 001	0 004	

Table B.15: Original and response verification bits per minute. Verification (q) indicates the bits per minute of the response verification trial itself. Theoretical best indicates the best bit rate that could have been achieved if all selected goals were kept and all selected non-goals were discarded during response verification. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0 68

Subject	Original	Keep Response if P3 was found in channel:				Keep Response if P3 was Absent in channel:				Verification (q) if P3 was Found in channel				Verification (q) if P3 was Absent in channel				Theoretical Best
		Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	Fz	Cz	Pz	Any	
1	14 45	14 45	14.45	14.71	14.98	13 44	13.44	12 71	12.48	6 42	6.42	7 11	8 17	3 15	3 15	1 41	0.97	14 71
2	14 23	14.23	14.52	13.95	15.40	13 95	14 23	14.52	14.23	2 64	3.22	2.10	5 30	2 10	2 64	3 22	2 64	11 23
3	17 66	17 66	17.66	17 66	19.81	17 41	17.66	17.66	17.41	10 41	10.41	10 41	18 31	9 61	10 41	10 41	9 61	17 66
4	14 45	14 45	14.99	14 19	15.27	14 19	14.19	14.19	14 19	3 49	5.11	2 78	6 02	2 78	2 78	2 78	2 78	11 15
5	9 09	9 09	9.09	9.09	13.10	9 09	9.09	9.28	9 28	0 14	0.14	0 14	11 17	0 14	0 14	0 02	0 02	9 09
6	14 62	11 99	14 62	14.62	15.74	14 26	14 26	15.74	14.99	2 84	1.24	1 21	7 79	0 26	0 26	7 79	2 84	15 36
7	12 37	12 62	11.17	11 88	13 91	11 64	12.37	12 37	11.64	2 71	0 09	0 99	7 27	0 59	2 05	2 05	0 59	13 65
8	12 92	12 02	12.46	12 92	12 69	12 92	12 92	12 92	12.92	1 00	2 22	3 86	2 99	3 86	3 86	3 86	3 86	13 15
9	11 63	11 63	11 63	11 63	11.88	11 63	11 63	11.63	11 63	0 22	0 22	0 22	0 81	0 22	0 22	0 22	0 22	11 63
Mean	13 49	13 16	13 40	13 41	14.75	13 17	13.31	13.45	13 20	3 32	3 23	3 24	7 58	2 52	2 84	3 53	2 62	13 77
Std	2 38	2 44	2.55	2 41	2 32	2 30	2 32	2 43	2 34	3 28	3.51	3 51	5 05	3 01	3 16	3 17	2 95	2 10

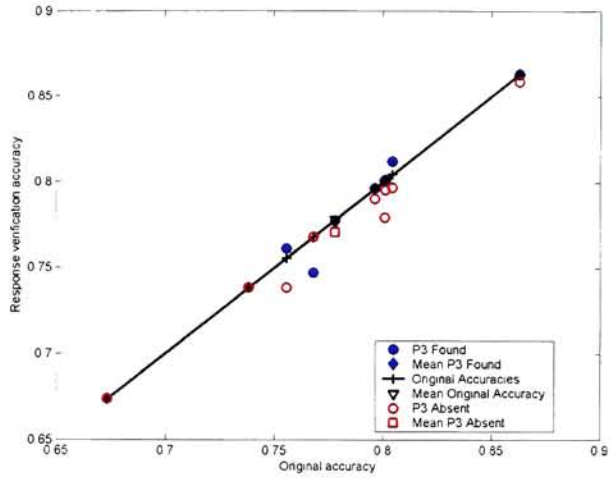


Figure B.32: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Fz). Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

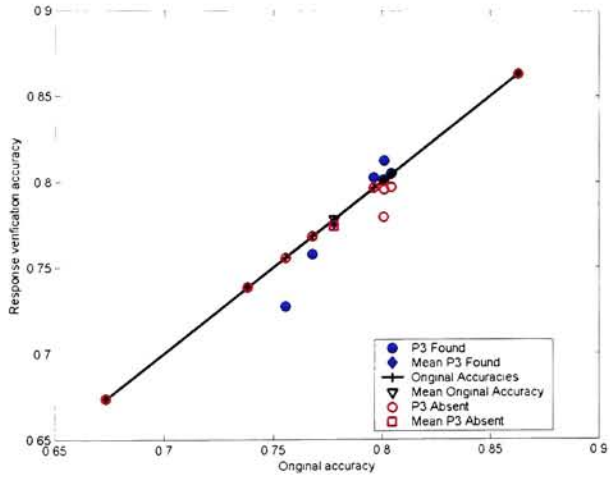


Figure B.33: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Cz). Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

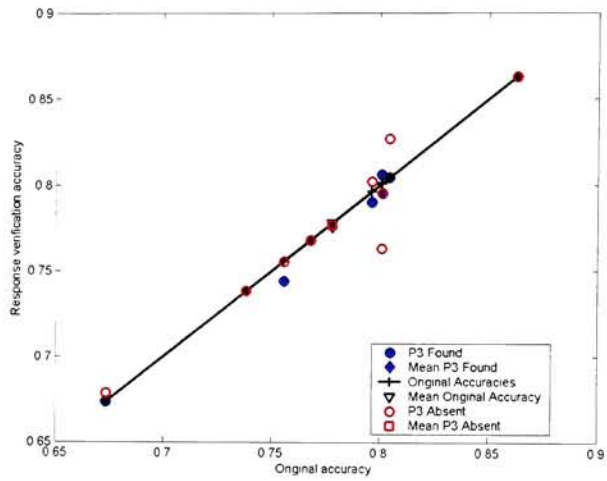


Figure B.34: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Pz). Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

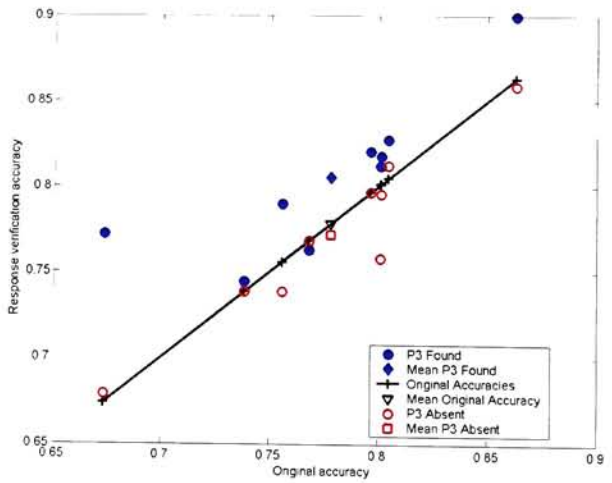


Figure B.35: Response Verification Accuracy versus Original Accuracy for P3 Found and P3 Absent methods (site Any). Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68.

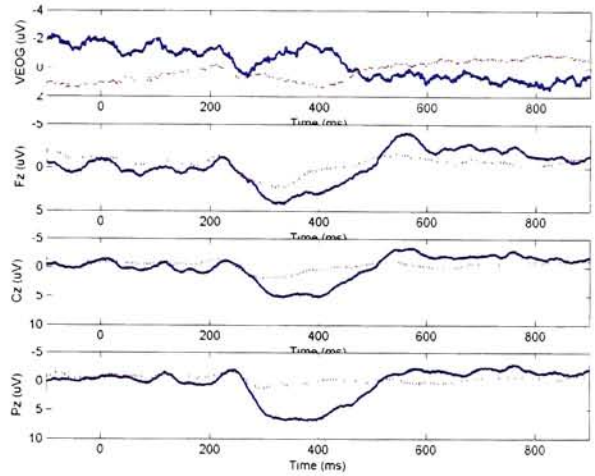


Figure B.36: Across All Subjects Response Verification Trials Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

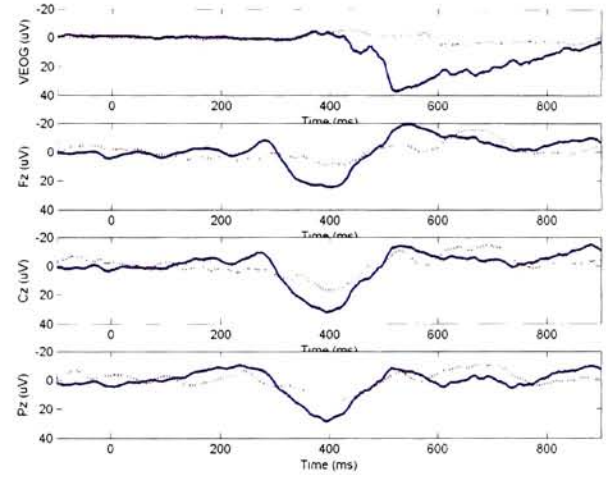


Figure B.37: Response Verification Trials Subject 1 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

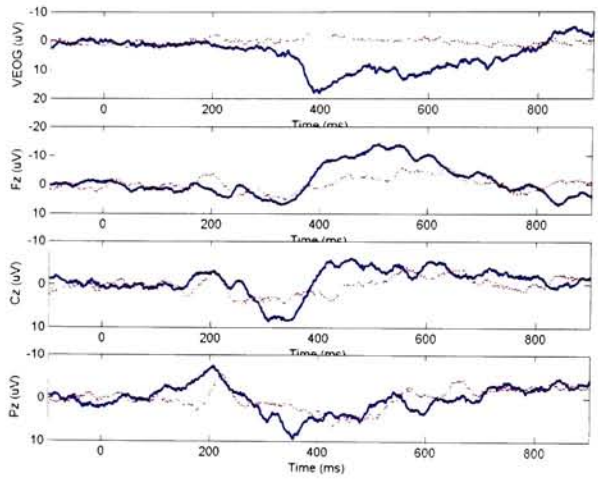


Figure B.38: Response Verification Trials Subject 2 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

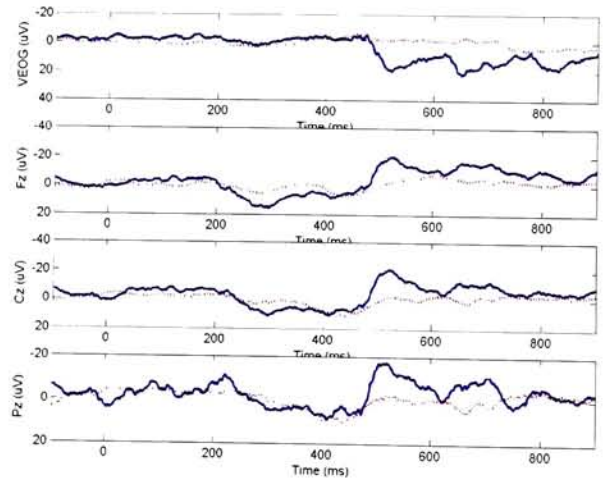


Figure B.39: Response Verification Trials Subject 3 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

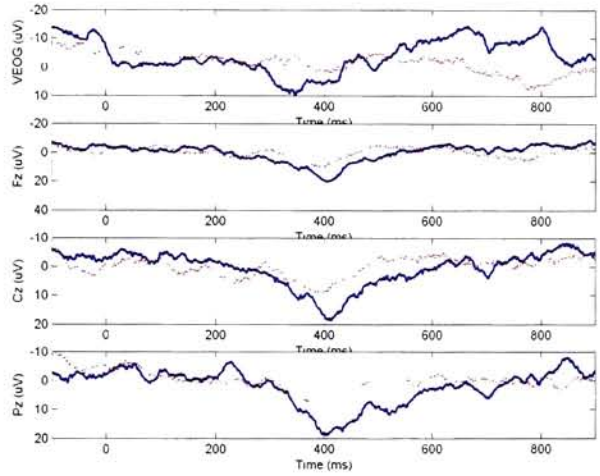


Figure B.40: Response Verification Trials Subject 4 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

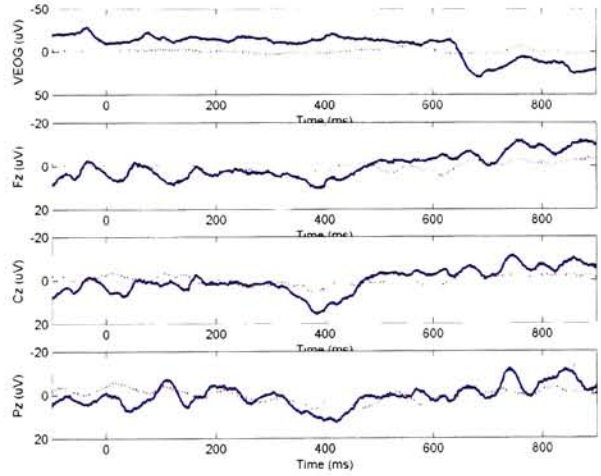


Figure B.41: Response Verification Trials Subject 5 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

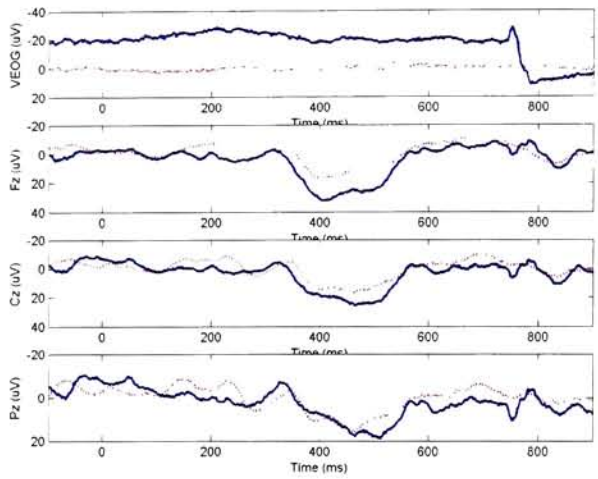


Figure B.42: Response Verification Trials Subject 6 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

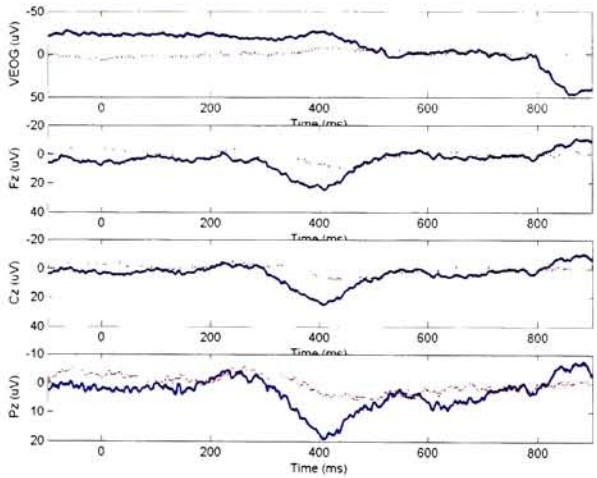


Figure B.43: Response Verification Trials Subject 7 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

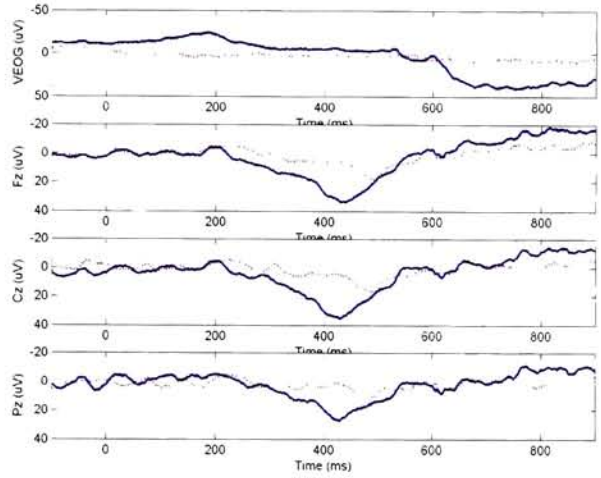


Figure B.44: Response Verification Trials Subject 8 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

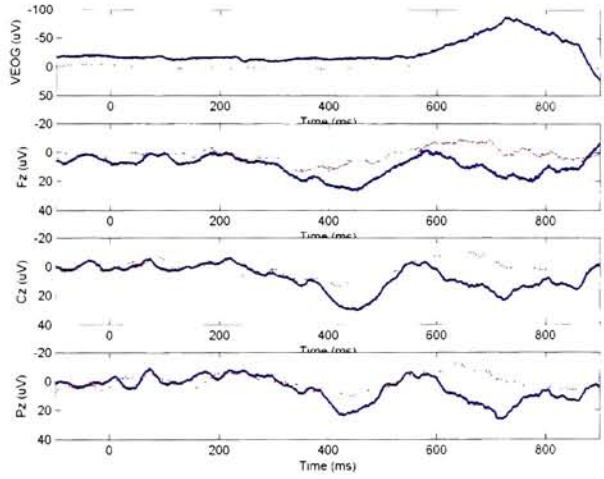


Figure B.45: Response Verification Trials Subject 9 Goal and Non-Goal Grand Averages. Keep response if P3 was found at site Any. Condition: Monitor, Algorithm: Correlation Indeterminate, Regressed: yes, EOG Threshold: 90, Correlation Threshold: 0.68

Bibliography

[1] http://www.eegspectrum.com, April 2003.

[2] Cyberlink. http://www.brainfingers.com/, April 2003.

[3] Opiates act on many places in the brain and nervous system. http://teens.drugabuse.gov/mom/mom_opi5.asp, April 2003.

[4] Dana H. Ballard. *An Introduction to Natural Computation.* The MIT Press, 2000.

[5] J.D. Bayliss. The use of the P3 evoked potential component for control in a virtual apartment. *IEEE Transactions on Rehabilitation Engineering*, 11(2). 2003.

[6] J.D. Bayliss and D.H. Ballard. Recognizing evoked potentials in a virtual environment. In S.A. Solla, T.K. Leen, and K. Müller, editors, *Advances in Neural Information Processing Systems 12*, Cambridge, MA, 2000. MIT Press.

[7] Jessica D. Bayliss. *A Flexible Brain-Computer Interface.* PhD thesis, University of Rochester, 2001.

[8] Jennifer Y. Bennington and John Polich. Comparison of P300 from passive and active tasks for auditory and visual stimuli. *International Journal of Psychophysiology*, 34(2):171–177, 11 1999.

[9] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398(6725):297–298, 1999.

[10] Niels Birbaumer, Thilo Hinterberger, Andrea Kübler, and Nicola Neumann. The thought-translation device (ttd): Neurobehavioral mechanisms and clinical outcome. *IEEE Transactions on Neural Systems and Rehabilitation Engineering,*, 11(2):120–123, 2003.

[11] Niels Birbaumer, Andrea Kübler, Nimr Ghanayim, Thilo Hinterberger, Jouri Perelmouter, Jochen Kaiser, Iver Iversen, Boris Kotchoubey, Nicola

Neumann, and Herta Flor. The thought translation device (ttd) for completely paralyzed patients. *IEEE Transactions on Rehabilitation Engineering*, 8(2):190–193, 2000.

[12] G. E. Birch and S. G. Mason. Brain-computer interface research at the neil squire foundation. *IEEE Transactions on Rehabilitation Engineering*, 8(2):193–195, 2000.

[13] G. E. Birch, S. G. Mason, and J. F. Borisoff. Current trends in brain-computer interface research at the neil squire foundation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), June 2003.

[14] B. Blankertz, G. Dornhege, C. Schäfer, R. Krepki, J. Kohlmorgen, K. R. Müller, V Kunmann, F. Losch, and G. Curio. Boosting bit rates and error detection for the classification of fast-paced motor commands based on single-trial eeg analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):100–104, 2003.

[15] B. Blankertz, C. Schäfer, G. Dornhege, and G. Curio. Single trial detection of eeg error potentials: A tool for increasing bcitransmission rates. pages 1137–1143. Artificial Neural Networks - ICANN 2002, 2002.

[16] R.M. Chapman and H.R. Bragdon. Evoked responses to numerical and non-numerical visual stimuli while problem solving. *Nature*, 203:1155–1157, 1964.

[17] Rémi Coulom. Simple chess engine. http://remi.coulom.free.fr/, August 2003.

[18] Jeffrey R. Cram, Glenn S. Kasman, and Jonathan Holtz. *Introduction to Surface Electromyography*. Aspen Publishers, Inc., 1998.

[19] Ct and spect scans, cited in [27]. `http://agora.leeds.ac.uk/comir/research/brains/brains.html`, September 2000.

[20] José del R. Millán and Josep Mouriño. Asynchronous bci and local neural classifiers: An overview of the adaptive brain interface project. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):159–161, 2003.

[21] Eamon Doherty, Gilbert Cockton, Chris Bloor, and Dennis Benigno. Improving the performance of the cyberlink mental interface with yes / no program. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 69–76. ACM Press, 2001.

[22] Eamon Doherty, Gary Stephenson, and Walter Engel. Using a cyberlink mental interface for relaxation and controlling a robot. *SIGCAPH Comput. Phys. Handicap.*, (68):4–9, 2000.

[23] Eamon P. Doherty, Gilbert Cockton, Chris Bloor, Joann Rizzoc, Bruce Blondina, and Bruce Davis. Yes/no or maybe further evaluation of an interface for brain-injured individuals. *Interacting with Computers*, 14(4), July 2002.

[24] Emanuel Donchin, Kevin M. Spencer, and Ranjith Wijesinghe. The mental posthesis: Assessing the speed of a P300-Based Brain-Computer Interface. *IEEE Transactions on Rehabilitation Engineering*, 8(2):174–179, 2000.

[25] L.A. Farwell and E. Donchin. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroenceph. Clin. Neurophysiol.*, pages 510–523, 1988.

[26] Bruce J. Fisch. *Fisch and Spehlmann's EEG Primer*. Elsevier Science BV, 3 edition, 1999.

[27] Simon Fisk. Music mind & matter. `http://www-users.york.ac.uk/~scf104/neuralmusic/overview.html`, October 2002.

[28] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller. How many people are able to operate an eeg-based brain-computer interface (bci)? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):145–147, 2003.

[29] Patricia F. Harner and Theda Sannit. *A Review of the International Ten-Twenty System of Electrode Placement*. Grass Instrument Company, 1974.

[30] Sara Ilstedt Hjelm and Carolina Browall. Brainball using brain activity for cool competition. *NordiCHI 2000 proceedings*, 2000.

[31] David A. Huffman. A method for the construction of minimum-redundency codes. *Proceedings of the I.R.E.*, 40:1098–1101, 1952.

[32] H.H. Jasper. The ten-twenty electrode system of the international federation. *Electroencephalogram and Clinical Neurophysiology*, 10:371–375, 1958.

[33] Keith A. Johnson and J. Alex Becker. Whole brain atlas. `http://www.med.harvard.edu/AANLIB/cases/caseM/mr1/035.html`, October 2002.

[34] T.P. Jung, C. Humphries, T. Lee, S. Makeig, M.J. McKeown, V. Iragui, and T.J. Sejnowski. Extended ica removes artifacts from electroencephalographic recordings. In M.I. Jordan, M.J. Kearns, and Sara A. Solla, editors, *Advances in Neural Information Processing Systems 10*, Cambridge, MA, 1998. MIT Press.

[35] P. R. Kennedy, R. A. E. Bakay, M. M. Moore, K. Adams, and J. Goldwaithe. Direct control of a computer from the human central nervous system. *IEEE Transactions on Rehabilitation Engineering*, 8(2):198–202, 2000.

[36] E. Lalor, S.P. Kelly, C. Finucane, R. Smith, R. Burke, R. B. Reilly, and G. McDarby. Steady-state vep-based brain computer interface control in an immersive 3-d gaming environment. *Journal of Applied Digital Signal Processing Trends of Brain Computer Interfaces Special Issue.* In submission.

[37] Janne Lehtonen. Eeg-based brain computer interface. Master's thesis, Helsinki University of Technology, April 2002.

[38] Carlo J. De Luca. Surface electromyography: Detection and recording. Unpublished work, 2002.

[39] Michael A. Paradiso Mark F. Bear, Barry W. Connors. *Neuroscience: Exploring the Brain.* Lippincott Williams and Wilkins, 2 edition, 2001.

[40] S. G. Mason and G. E. Birch. A general framework for brain-computer interface design. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(1):70–85, 2003.

[41] J. R. Wolpaw D. J. McFarland and T. M. Vaughan. Brain-computer interface research at the wadsworth center. *IEEE Transactions on Rehabilitation Engineering*, 8(2):222–226, 2000.

[42] P. Meinicke, M. Kaper, F. Hoppe, M. Heumann, and H. Ritter. Improving transfer rates in brain computer interfacing: a case study. In S. Becker, S. Thrun, and K. Obermayer, editors, *Advances in Neural Information Processing Systems 15*, Cambridge, MA, 2003. MIT Press.

[43] Stefan Christ Michael Falkenstein, Jörg Hoormann and Joachim Hohnsbein. Erp components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, 51(2-3):87–107, January 2000.

[44] M. Middendorf, G. McMillan, G. Calhoun, and K. S. Jones. Brain-computer interfaces based on the steady-state visual-evoked response. *IEEE Transactions Rehabilitation Engineering*, 8:211–214, 2000.

[45] Karl E. Misulis and Toufic Fakhoury. *Spehlmann's Evoked Potential Primer.* Butterworth-Heinemann, 3 edition, 2001.

[46] Ernst Niedermeyer. *Electroencephalography: basic principles, clinical applications, and related fileds.* Lippincott Williams & Wilkins, 4 edition, 1999.

[47] S. Nieuwenhuis, K. Ridderinkhof, J. Blom, G. Band, and A. Kok. Error-related brain potentials are diferentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology*, 38:752–760, 2001.

[48] Bernhard Obermaier, Gernot Müller, and Gert Pfurtscheller. 'virtual keyboard' controlled by spontaneous eeg activity. In *LNCS 2130*, pages 636–641, 2001.

[49] Introduction to fmri. `http://www.fmrib.ox.ac.uk/fmri_intro/fusion.gif`, October 2002.

[50] L.C. Parra, C.D. Spence, A.D. Gerson, and P. Sajda. Response error correction - a demonstration of improved human-machine performance using real-time eeg monitoring. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):173–177, 2003.

[51] Lucas Parra, Chris Alvino, Akaysha Tang, Barak Pearlmutter, Nick Yeung, Allen Osman, and Paul Sajda. Linear spatial integration for single-trial detection in encephalography. *NeuroImage*, 17:223–230, 2002.

[52] J. Perelmouter and N. Birbaumer. A binary spelling interface with random errors. *IEEE Transactions on Rehabilitation Engineering*, 8(2):227–232, 2000.

[53] J. Perelmouter, B. Kotchouby, A. Kübler, E. Taub, and N. Birbaumer. Language support program for thought-translation devices. *Automedica*, 18:67–84, 1999.

[54] G. Pfurtscheller. Functional brain imaging based on erd/ers. *Vision Research*, 41(10-11):1257–1260, 2001.

[55] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper. Brain oscillations control hand orthosis in a tetraplegic. *Neuroscience Letters*, 292(3):211–214, 2000.

[56] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlögl, B. Obermaier, and M. Pregenzer. Current trends in graz brain-computer interface (bci) research. *IEEE Transactions on Rehabilitation Engineering*, 8(2):215–219, 2000.

[57] G. Pfurtscheller, C. Neuper, G. R. Müller, B. Obermaier, G. Krausz, A. Schlögl, R. Scherer, B. Graimann, C. Keinrath, D. Skliris, M. Wrtz, G. Supp, and C. Schrank. Graz-bci: State of the art and clinical applications. *IEEE Transactions on Neural Systems and rehabilitation Engineering*, 11(2):177–180, 2003.

[58] G. Pfurtscheller, Ch. Neuper, C. Andrew, and G. Edlinger. Foot and hand area mu rhythms. *International Journal of Psychophysiology*, 26(1-3):121–135, 1997.

[59] G. Pfurtscheller, Ch. Neuper, D. Flotzinger, and M. Pregenzer. Eeg-based discrimination between imagination of right and left hand movement. *Electroencephalography and Clinical Neurophysiology*, 103(6):642–651, 1997.

[60] Gert Pfurtscheller. *Electroencephalography: basic principles, clinical applications, and related fileds.* Lippincott Williams & Wilkins, 4 edition, 1999.

[61] Gert Pfurtscheller, Gernot R. Müller, Jorg Pfurtscheller, Hans Jurgen Gerner, and Rudiger Rupp. 'thought' - control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia. *Neuroscience Letters*, 351(1):33–36, 2003.

[62] Jaime A. Pineda, David S. Silverman, Andrey Vankov, and John Hestenes. Learning to control brain rhythms: Making a brain-computer interface possible. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):181–184, 2003.

[63] James B. Polikoff, H. Timothy Bunnell, and Winslow J. Borkowski Jr. Toward a P300-based Computer Interface. `http://www.asel.udel.edu/speech/reports/resna95/p300.pdf`, 1996.

[64] Edward L. Reilly. *Electroencephalography: basic principles, clinical applications, and related fileds.* Lippincott Williams & Wilkins, 4 edition, 1999.

[65] Gerhard Roth. The quest to find consciousness. *Scientific American Mind*, 14(1):32–39, 2004.

[66] A. James Rowan and Eugene Tolunsky. *Primer of EEG.* Elsevier Science (USA), 2003.

[67] Chang Su Ryu, Yoonseon Song, Done-Sik Yoo, Sangsup Choi, Sung Sill Moon, and Jin-Hun Sohn. Eeg-based discrimination between yes and no. *Proceedings of the First Joint BMES/EMBS Conference Serving Humanity, Advancing Technology*, page 444, 1999.

[68] Renato M. E. Sabbatini. Mapping the brain. `http://www.epub.org.br/cm/n03/tecnologia/eeg.htm`, September 1997. Last accessed October 27, 2002.

[69] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw. Bci2000: A general-purpose brain-computer interface (bci) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–1043, 2004.

[70] Gerwin Schalk, Jonathan S Carp, and Jonathan R Wolpaw. Temporal transformation of multiunit activity improves identification of single motor units. *Journal of Neuroscience Methods*, 114:87–98, 2002.

[71] Gerwin Schalk, Jonathan R. Wolpaw, Dennis J. McFarland, and Gert Pfurtscheller. Eeg-based communication: presence of an error potential. *Clinical Neurophysiology*, 111(12):2138–2144, 2000.

[72] H.V Semlitsch, P. Anderer, P Schuster, and O. Presslich. A solution for reliable and valid reduction of ocular artifacts applied to the P300 erp. *Psychophys.*, 23:695–703, 1986.

[73] Jim Strommer. Tutorial: Clinical pet neurology. `http://www.crump.ucla.edu/software/lpp/clinpetneuro/function.html`, October 2002.

[74] Patrick Suppes, Zhong-Lin Lu, and Bing Han. Brain wave recognition of words. *Proc. Natl. Acad. Sci.*, 94:14965–14969, 1997.

[75] S. Sutton, M. Braren, J. Zublin, and E. John. Evoked potential correlates of stimulus uncertainty. *Science*, 150:1187–1188, 1965.

[76] Aaron Tay. Chess engines cutting through the confusion. http://www.chesskit.com/aarontay/Winboard/confusion.html, October 2002.

[77] Brain facts and figures. `http://faculty.washington.edu/chudler/facts.html`, November 2002.

[78] F. Vidal, B. Burle, M. Bonnet, J. Grapperon, and T. Hasbroucq. Error negativity on correct trials: a reexamination of available data. *Biological Psychology*, 64(3):265–282, November 2003.

[79] J. J. Vidal. Towards direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2:157–180, 1973.

[80] J. J. Vidal. Real-time detection of brain events in eeg. *IEEE Proc*, 65:157–180, 1977.

[81] Harm Johannes Wieringa. Meg, eeg and the integration with magnetic resonance images. `http://www.neuro.com/megeeg/index2.htm`, 1992. Last accessed October 27, 2002.

[82] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller. Eeg-based communication: improved accuracy by response verification. *IEEE Transactions on Rehabilitation Engineering*, 6(3):326–333, 1998.

[83] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, and Theresa M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.

[84] J.R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J.McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A.Quatrano, C. J. Robinson, and T. M. Vaughan. Brain-computer interface technology: a review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, 8(2):164–173, 2000.