4-18-2014

Can Portable EEG Headsets be Used to Determine if Students are Learning?

Wesley Adam Boyd

CAN PORTABLE EEG HEADSETS BE USED TO DETERMINE IF STUDENTS ARE LEARNING?

by

Wesley Adam Boyd

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Major: Psychology

The University of Memphis

May 2014

Dedication

This thesis is dedicated to my son, Bradley. Since he was only eight months old, I have raised him as a single father while pursuing a college education at the University of Memphis. I hope this serves as evidence to him that, even through the toughest challenges in life, you can accomplish anything you set out to do as long as you try.

Acknowledgements

I would like to acknowledge the following people who have been paramount to my success as a graduate student:

First and foremost, I would like to thank God for giving me the tenacity and patience to complete this project. I would like to thank my advisor, mentor, and friend, Dr. Philip Pavlik, Jr. for all his guidance and for letting me stake the claim of being the first graduate student to complete a thesis under his mentorship. I would like to thank the Optimal Learning Lab and its members for all of their help and support. Special thanks go out to my committee members and mentors, Dr. Jeff Sable, Dr. Roger Kreuz, and Dr. Gayle Beck for all of their support and guidance. I would also like to thank my other academic mentors Dr. Charles Blaha, Dr. Xiangen Hu, and Dr. Benjamin Nye. Thank you to all of the members of IIS and all of the Psychology Department at the University of Memphis. Many thanks to my Mom, Lisa Boyd, and my Dad, Johnny Boyd for everything they do. Without them, this would not have been possible. Thanks to my Granny, Margaret Jones, and to all my family for believing in me. Thanks to my son, Bradley, and to my fiancée, Jenna, who endured the late nights of studying and all the EEG practice talks I gave them. Thank you to all my friends who motivate and continue to inspire me.

Abstract

Boyd, Wesley Adam. M.S. The University of Memphis. May 2014. Can Portable EEG Headsets be Used to Determine if Students are Learning? Major Professor: Philip Pavlik, Jr., Ph.D.

This study examined EEGs recorded from a single channel, portable EEG headset (NeuroSky MindWave) during the study period of a paired-associate word paradigm which used Swahili words and their English meanings. It was hypothesized that there would be a significant difference in gamma, theta, and beta band powers for when students recalled words correctly vs. when they did not recall correctly on a subsequent test. There were 35 participants who consisted of students that volunteered at the University of Memphis (20 females and 15 males, 31 of which were right-handed and 4 which were left-handed). A paired-samples *t*-test suggested that there was a higher mean z-score for brainwave activity during the study period in the high gamma range (41 - 49.75Hz) for when participants did not recall words correctly on a test, which was opposite of what previous research has found regarding encoding. Based on the results of this study, the MindWave seems to capture muscle activity and/or saccadic behavior that is suggested by higher gamma maximums on average in the study period for word-pairs which resulted in failed recall. Exploratory results may lend insight to future work using portable EEG devices. This study's main objective was to determine if portable EEG devices could be used to determine when students learn new information. Further testing, especially using other portable EEG devices is necessary to answer this question.

TABLE OF CONTENTS

Appendices

Can Portable EEG Headsets be Used to Determine if Students are Learning?

Introduction

In past research, many types of physiological measures, such as biosensors used to determine anxiety level, have been used to collect and integrate data into computer-tutoring software so that programs can adapt to students and promote high levels of involvement while interacting with computer agents (Rani, Sarkar, & Liu, 2005). While devices like biosensors may be beneficial, the practicality and use of such equipment is not feasible in a real-world setting where equipment needs to be relatively inexpensive, portable, and easy to use. However, a few researchers (e.g., Chang, Nelson, Pant, & Mostow 2013; Choi, Jones, & Schwartz, 2012) have explored the use of portable electroencephalography (EEG) devices in real-world settings such as the workplace and educational settings. Using a portable EEG device in a classroom could aid in providing customized, real-time feedback to the instructor via computer software and could also provide feedback to an intelligent tutoring system (ITS). By researching students' brainwave activity and how it might indicate certain cognitive functions, such as memory encoding during study, progress can be made in moving these valuable tools into the classroom in an attempt to enhance student learning.

EEG is the recording of electrical activity from the scalp, which was first recorded in a human by Hans Berger in 1924 (Berger, 1929). EEG is a non-invasive technique that records electrical activity from wherever pairs of metal electrodes from an EEG device are placed on the scalp. This electrical activity is caused by the activation of neurons which results in the flow of currents. EEG measures changes in electrical potentials at the scalp that are caused by postsynaptic graded potentials from pyramidal cells (Teplan, 2002). Generally speaking, the summation of this electrical activity of neurons firing in synchrony is what produces the EEG data. Due to these currents being very weak, the EEG device amplifies the signal and then outputs the raw signals. Portable EEG devices use a fast-fourier transform (FFT) to process the raw signals to output the frequency bands.

Brainwaves occur at different frequencies and amplitudes. These brainwave frequencies are usually grouped into five different bands (Niedermeyer & Lopes da Silva, 2004). Each frequency range correlates with characteristic effects seen in each range (e.g., cognitive experiences like concentration and alertness). The five frequencies are the Delta (0.1-3.5 Hz), Theta (4-7.5 Hz), Alpha (8-13 Hz), Beta (14-30 Hz), and Gamma frequencies (30-100Hz). Delta waves are the highest in amplitude and are seen in deep sleep. Theta waves are seen in light sleep states (including dream sleep) and have also been seen in states of meditation, deep relaxation, and drowsiness. Alpha waves are seen in states of calmness, relaxation, and reflection. Beta waves are seen in the normal, active awake state and are associated with attention, alertness, and engaging activity. Gamma waves are seen in higher learning, memory processing, and have been linked to Tibetan Buddhist monk meditation (O'Nuallain, 2009).

The synchronous neural firings recorded by an EEG device have been found to correlate highly with cognitive processing. Brain synchrony is defined as the simultaneous in-phase activity of many neurons that are spread over a certain area, much like the well-known military cadence in which many soldiers march and sing together. These soldiers are in synchrony. The more soldiers there are, the louder they march and sing. Similarly, strength increases as more synchronous neuronal activity occurs.

In the past, recording brainwave activity was typically performed in a lab setting where many electrodes were used along with a converter box and other cumbersome peripherals. Also, typical EEG setups require cleaning each electrode site along with application of saline gel which is very time-consuming. Now, the portable EEG headset creates research opportunities in real-world settings where they were previously not feasible (e.g., in the classroom or in a moving automobile). If brainwave data can be indicative of learning, then it could be used as a type of software feedback to enhance learning in tutoring sessions. For instance, previous research found that during a review period (before a subsequent memory recall test), a gamma band increase during encoding correlated with recall at a later test (Gruber, Tsivilis, Montaldi, & Muller, 2004). Replication of these findings using portable EEG devices would help brainwave data move into real-world applications.

Literature Review

To understand the possible practical utility of portable EEG, we can draw upon the many laboratory studies over the past few decades which showed how cognitive processes (like memory processing) related to particular EEG frequencies. Studies have been conducted looking at gamma band activity's role in cognitive processing. Gamma band activity has been shown to be associated with learning and memory processing. For example, Hermann, Lenz, Junge, Busch, and Maess (2004) found that gamma activity is evoked by visual stimuli that match existing representations in long term memory (LTM). This memory comparison is said to elicit gamma activity for other stimuli as well, such as known objects (spoon, fork, car) vs. unknown non-objects (shapes that are not easily distinguishable). Conscious access to both short term memory (STM) and LTM is said to modulate gamma activity. More support for this claim is shown in that words evoke more gamma activity than pseudo-words, which is a result of these words being represented in our mental lexicon or LTM (Pulvermüller, Lutzenberger, Preissl, & Birbaumer, 1995). Some studies have found that significant gamma power increases happen when certain cognitive processes are experienced. For instance, amplitude increases in the gamma band have been observed when a person becomes more attentive (Bouyer, Montarom, Vahnee, Albert, & Rougeul, 1987). Also, increases in frontal gamma activity (around 40 Hz) were found in those who selectively attended to stimuli as opposed to those who ignored stimuli (Tiitinen et al., 1993). This may show that selectively attending to auditory, somatosensory, or visual stimuli (Chen & Herrmann, 2001; Hoogenboom et al., 2006; Pantev, 1995) elicits gamma band activity and thus the EEG may aid in detecting whether or not a person is paying attention.

In a study where participants were instructed to attend to specific targets among different sets of visual stimuli, the stimulus sets were thought to undergo comparisons with existing representations of the target template in STM—that is to say, the participants were thought to be unconsciously comparing the non-target stimuli with the intended targets, which were already represented in their memories (Hermann, Mecklinger, & Pfeifer, 1999; Hermann et al., 2004). Interestingly, what seemed to control the strength of the gamma band response was the similarity of features that non-targets shared with the intended target. These findings lend support to the idea that gamma band activity is induced when our brains make comparisons based on existing information (called memory matching), either in LTM or STM. This may be helpful in learning by enhancing the ability to detect generalization and discrimination based on gamma band power during a learning session. Associations between gamma band activity and behavioral performance have been said to hold direct relationships with memory processes, where gamma band responses occur during the encoding of memories and predict subsequent recall (Hermann, Frund, & Lenz, 2010). In light of these findings, an increase in gamma band activity during encoding could be expected if a person successfully encodes a memory.

Many other studies have shown the association between gamma band activity during memory encoding and subsequent memory performance. In a study where epileptic patients studied wordlists during a memory paradigm, subsequent recall of those words was predicted by increases in gamma band power during the encoding period (Sederberg, Kahana, Howard, Donner, & Madsen, 2003). Similarly, in a study where participants may successfully encode new words and subsequently be tested on those words, an increase in gamma band activity in the encoding period might be possible for words remembered as opposed to words that are forgotten.

In addition to gamma band oscillations being predictive in memory tasks, theta oscillations have also been shown to be associated with memory processes (Klimesch et al., 2001). By using a remember/know paradigm (used to effectively separate these two distinct forms of awareness during retrieval), Klimesch and colleagues (2001) investigated the associations between the conscious experiences of remembering and knowing and theta band activity during these events. Participants in this study were first presented with a set of words for a short duration and then were instructed to, in the next phase, make judgments based on their recollection of the words—that is, whether they remembered seeing a word in the list, knew the words were in the list, or did not recollect seeing the word and therefore determined it to be a “new” word. The EEGs of the participants were recorded throughout the experiment. An increase in theta activity was found in all three correct judgments (remember, know, and new). However, the duration and onset of each was different. The new word judgments showed the shortest duration of theta increases while the remembered word judgments showed the longest duration. Onset of synchronous theta increases was first apparent in knowing (300-450 ms) and then in remembering (450-625 ms). Klimesch and colleagues (2001) postulated that this was evidence of two distinct neural correlates for remembering and knowing in memory. In a study that found gamma and theta band oscillations to be predictive of the encoding and retrieval of declarative memory, Osipova et al. (2006) suggested that theta oscillations play a role in synaptic plasticity that facilitates memory encoding. Osipova and colleagues (2006) extended upon Klimesch’s “subsequent memory effect” (remembering an encoded item on a test) findings by using pictorial stimuli instead of words. These increases in gamma and theta power were said to occur within 0.3-1s of onset in the encoding period, which indicated a momentary binding of memory. Gamma and theta power increases were found to be associated with memory processing for successful encoding and retrieval, and have also been shown to be evident in the maintenance of information in working memory (Jensen & Tesche, 2002). These findings are consistent across researchers and have been replicated numerous times. Extending these findings using portable EEG could aid in answering the question as to whether these portable devices are capable of being used reliably in applied settings.

Other frequency bands, such as the alpha and beta bands, have also been found to be associated with cognitive processing. Further examination of these bands’ activities is needed to determine their role in possible memory processing. Let us consider some of the recent findings that might help to show their importance in this study.

Changes in beta band activity have also been proposed to be associated with paying attention (Egner & Gruzelier, 2004). Training in the low beta frequency range (15-18 Hz; called beta1 training) and the sensorimotor rhythm (12-15 Hz) was shown to increase reaction times in certain attention tasks and indicate a general attention-enhancing effect. Egner and Gruzelier (2004) postulated that this was a result of increased arousal in a noradrenergic alertness network of attention. This suggests that training-evoked low beta activity (12-15 Hz) enhances attention and could be seen by an increase in power in the beta band. Interestingly, beta band activity was also found to be modulated by attention, which has been found to be directly related to associative learning (Asaad, Ranier, & Miller, 1998; Bardouille, Picton, & Ross, 2010; Kruschke, 2001). If attention modulates beta activity, then it is plausible to suggest that when a person is paying attention in an associative learning scenario, beta power will increase if learning has taken place.

Alpha band activity normally accompanies cognitive processes such as calmness and relaxation. However, contrary to the classical theory in which alpha band frequencies represent "brain idling," alpha band oscillations during a short term memory task have been found to increase with memory load during the retention period (Jensen, Gelfand, Kounios, & Lisman, 2002). However, EEG effects have been thought to also depend on the type of task. For example, a study that examined EEGs, specifically the theta and alpha activities, used an *n* back task. Results of the experiment using an *n*-back task, a task in which subjects are presented with a continuous stream of items and are asked to indicate whether the displayed item matches the one *n* positions back, showed that theta activity increased with memory load and alpha activity decreased (Gevins, Smith, McEvoy, & Yu, 1997). Using a modified version of the Sternberg task (Sternberg, 1966), Jensen and colleagues were able to study the temporal and spatial development of alpha band activity during encoding, retention, and recognition, which were separated in time. The classic Sternberg task allows for this separation naturally because it involves an encoding period (presentation of items to memorize), a memory maintenance period (retention), and finally a judgment period which presents a "probe" item in which the subject must judge whether or not that probe item was in their memory list. EEGs were recorded during the retention and recognition intervals of the task. Analysis of the data showed that the alpha rhythm was the most salient frequency showing an increase in amplitude with memory load, an evident spectral peak, and very apparent oscillations. Power spectra were calculated for each trial and then averaged. The average spectral peak was ~11 Hz, and the area in which alpha activity was strongest was over the posterior region which was said to be indicative of parietal-occipital sulcus (a well-known alpha source) contributions.

The fact that amplitude increases with load suggests that the alpha band has a graded quality that depends on the number of items being stored in memory. Another important result that signifies a temporal relationship between working memory and alpha activity is that after the memory task was complete and subjects no longer needed to remember the items from the task, the alpha band power decreased within a few hundred milliseconds (Jensen et al., 2002). Jensen and his colleagues stated that the difference between the Sternberg task and the *n*-back task that Gevins et al. (1997) used might explain why the two studies found different results. They speculated that since the *n*-back task is very demanding (includes many overlapping operations), participants may have employed visual strategies which may have taxed the visual system under memory load and thus decreased alpha activity (given that alpha activity indicates a suppression of processing in visual areas; Fu et al., 2001). In the Sternberg task, operations are temporally separated, and using a visual strategy would be unlikely. These findings suggest that alpha band activity is associated with memory tasks in that alpha power has a graded increase during working memory so long as the visual system is not severely overworked.

Portable EEG

In recent years, portable EEG devices, such as the NeuroSky MindWave (MW) and NeuroSky Mindset (MS), have been used in research to investigate their potential as tools for revealing useful information about brainwave frequencies associated with mental states, such as attention and meditation. These lightweight headset devices feature single, dry electrodes (single-channel-referenced to the ear lobe) that do not require the use of a saline gel like the cumbersome and more complex traditional EEG equipment. The single channel measures electrical signals produced by neural activity in the frontal lobe region and converts those signals into useable binary data. These devices are often referred to as Brain Computer Interfaces (BCI). BCIs have been proposed to work well in learning environments where biofeedback could be used to enhance learning (Chang et al., 2013).

NeuroSky's devices record raw brainwave data and outputs the powers and raw signals of the delta, theta, alpha (high and low), beta (high and low), and gamma (high and low) frequencies(see Appendix C for ranges). It also outputs two custom measures, "attention" meter values, which can indicate the level of mental focus, and "meditation" meter values, which can indicate the level of mental calmness. These are referred to as NeuroSky's eSense meters. NeuroSky describes the computations to determine these values as trade-secrets that cannot be divulged. Nevertheless, research has shown that these two measures are able to clearly and reliably differentiate between higher states of cognition and emotional response (Crowley, Sliney, Pitt, & Murphy, 2010).

Rebelledo-Mendez and De Freitas (2008) assessed the MS for reliability and usability. An assessment exercise was developed that used a model of attention which combined the learner's EEG inputs (MS readings) and the learner's actions in a computer-based learning environment. Furthermore, MS readings were combined with user-generated data to provide a more accurate representation of attention levels. These user-generated data included whether questions were answered correctly or incorrectly and the amount of time taken to answer each question. Rebelledo-Mendez and De Freitas's (2008) model of attention not only detects attention patterns, but also gives feedback to the participant. The model was deployed using a state-of-the-art AI-driven avatar (programmed in C#), which collects user-generated information as well as EEG data. The avatar asks multiple-choice questions while collecting data (EEG and user-generated) and transmits this data to the computer in which software communicates with the avatar. By using the combined data, the avatar is able to dynamically interact and adapt to each learner's performance behavior. The study ultimately showed a positive correlation between self-reported (questionnaire-based) and measured attention levels and the MS was shown to be reliable in providing accurate readings of attention. Results from analyzing the device's usability suggest that most users do not experience any significant problems due to head size. Furthermore, it was determined that when the device fit properly, it did provide valid and reliable data (Rebolledo-Mendez, et al., 2009). These findings validate the choice of using the single-channel EEG devices for research in an educational setting.

The NeuroSky MS, MW, and Myndplay (another variation of the device that features an elastic band) were tested against each other in an attempt to reveal the most reliable device when analyzing EEG data for safety improvement among industry professionals (Choi et al., 2012). Two mental states, alertness and drowsiness, were assessed and compared using BCI2000 (an open source software) and Matlab. Significant differences were found between alertness and drowsiness. The drowsy state showed less alpha and beta wave activity than the alert state—which suggests that being alert elicits more alpha and beta wave activity. Of the three, the MW was said to have produced the most accurate and reliable data due to its technical and physical attributes. Additionally, in another recent study (Chang et al., 2013), it was determined that the NeuroSky MindWave, "…has the potential to detect mental states relevant to tutoring, such as comprehension, engagement, and learning" (p. 18). It was also stated that conducting further research using this device could lead to it being used in a tutorial setting that could adapt and respond to an individual's mental state and thus improve learning in an educational setting (Chang et al., 2013). Originally created for the gaming industry, these single-channel EEG devices, including the MW, have shown promising results in the scientific community as a reliable research tool. Being highly affordable and easy to use, this device makes a great candidate to use in educational learning scenarios, especially in intelligent tutoring systems that adapt to students based on the student's mental state feedback. One such ITS that could benefit from a constant stream of brainwave data is AutoTutor, which is an intelligent tutoring system that uses an animated conversational agent to interact with a student using natural language (Graesser, Wiemer-Hastings, K., Wiemer-Hastings, P., & Kreuz, & Tutoring Research Group, University of Memphis, 2000). Since intelligent tutoring systems constantly undergo progressive changes, the very affordable MW could be utilized in developing an even more adaptable tutoring system that is tailored to each individual student.

Hypotheses

Considering past research, this study will mainly focus on analyzing the gamma, theta, and beta bands for which we have three main hypotheses. These frequencies have all been identified as being associated with memory processing. Gamma power increases have been found in paired-associate learning (PAL) tasks which can predict subsequent recall (Hermann et al., 2010), and gamma responses have generally been found to be evoked by memory encoding (Sederberg et al., 2003). It is hypothesized that there will be a significant difference in gamma wave power during encoding for word-pairs that are remembered correctly vs. word-pairs that are not remembered correctly on a subsequent test. Theta power increases have also been seen in memory encoding and have been suggested to coincide with neural plasticity (Osipova et al., 2006). It is hypothesized that there will be a significant difference in theta wave power during encoding for word-pairs that are remembered correctly vs. word-pairs that are not remembered correctly on a subsequent test. Beta increases have been thought to be evoked by attentiveness (Egner & Gruzelier, 2004), and have also been found to modulate beta activity during associative learning (Asaad et al., 1998; Bardouille et al., 2010; Kruschke, 2001). Finally, it is hypothesized that there will be a significant difference in beta power in the encoding period for word-pairs that are remembered correctly vs. word-pairs that are not remembered correctly on a subsequent test.

Exploration of all frequency bands is necessary to further discover how they may be associated with memory processing while learning and to also help determine what bands may contribute to construction of an adaptive personal learner model using EEG. For example, alpha power increases have been seen in working memory tasks which show a graded quality, that is, power increase is contingent on how much information is being processed (Jensen et al., 2002). In this instance, alpha band information might indicate a heavy processing load which could be helpful in determining when a learner is bogged down with too much information. However, it is unclear whether this type of information is meaningful, and exploration would help us drive future research using BCIs.

Methods

Materials

For the present experiment, the NeuroSky MindWave was chosen for its low cost, availability, and portability. This non-invasive headset features a single electrode that touches the forehead approximately two inches above the eyebrow and records brainwave activity in the frontal lobe at Fp1 (according to the International 10-20 System). This device does not require the use of saline gel. The headset features an adjustable headband that can accommodate different head sizes and offers enough rigidity to fit over various hairstyles. The MW samples at 512 Hz (512 times per second). The recorded data undergoes a Fast Fourier Transform (FFT) which averages and filters the brainwave data into an output of 1 Hz in an Excel file (it logs all the EEG data once per second). The MW connects through NeuroSky's Thinkgear Connector driver, and a custom Java application that we formed using a JSON controller reads and records through this proprietary connection.

Task Selection

According to the cell assembly theory proposed by Donald Hebb (1949), learning is said to occur when cells assemble and fire in synchrony. Gamma band responses have been found to be associated with this phenomenon, and paired-associative learning (PAL) paradigms have been used to test whether gamma responses are associated with cell assembly and associative learning (Gruber, Keil, & Müller, 2001; Marshall, Helgadóttir, Mölle, & Born, 2006; Miltner, Braun, Arnold, Witte, & Taub, 1999). This type of learning is considered explicit and conscious learning as opposed to unconscious or unintentional learning. Given the evidence that this paradigm has been used many times and is considered reliable, we chose to use a paired-associate word task which presents 17 Swahili words (and three phrases—*Hello, Goodbye,* and *Thank You*) along with their English meanings. The word pairs were presented in a random sequence for each participant for both learning and testing. Unlike some tasks (like the Sternberg task), this task did not involve an explicit retention period (although there is an implicit period in which the items are retained). There was an encoding period in which the word pairs were presented, a short instructions screen that told participants what to do in the test (typed response), and a testing period.

The words were oriented in the middle of the screen, one on top of the other (Swahili word on top, English meaning below). This arrangement was considered optimal for this experiment because after analyzing pilot data, we found that our original arrangement of Swahili word on the left-of-middle and English word on the right-of-middle may have elicited a decrease in alpha band activity as a result of visual system taxation (Jensen et al., 2002). Originally, we were going to use 30 Swahili words, but after running pilot tests we found that using 30 words resulted in poor recall along with impoverished data which was likely a result of participant fatigue or inattentiveness. So, after deciding to decrease the number of Swahili items, the most difficult words (phonetically) were eliminated and 17 words remained. The challenge of selecting how many items to include involved careful planning of how to acquire a balance of recalled items vs. non-recalled items so that we have good data to analyze that represents both conditions of the outcome variable (recall vs. no recall). So, after further pilot testing using 17 words, we determined that this number of carefully selected items gave us this balance and was optimal for this particular experiment.

Two files were created during the experiment. One file contained the EEG data, and the other file contained the word list which recorded the exact times of when a particular word was studied, when it was tested on, and whether or not it was correctly recalled. These files were time-matched in that when a particular word was being studied, the EEG file times were synchronized with the word file times (within 50-100ms). This, of course, allowed us to know the exact ten second time period in which a word was being studied along with the exact EEG data in that ten second time period.

Participants

Thirty-five students from the University of Memphis voluntarily took part in this study and received monetary compensation. Participants' ages ranged from 18 to 72 years, and the mean age for this study sample was 26.86 (M = 26.86, SD = 9.87). Gender and handedness was recorded. There were 20 females and 15 males in the sample (31 of which were right-handed, and 4 were left-handed). Written informed consent was obtained from each participant, and all ethical regulations were followed according to the Declaration of Helsinki. The Experimentation Protocol was submitted to and approved by the University of Memphis's Institutional Review Board (IRB). The experimenter had the proper training and certification to conduct experiments using human subjects. Those who were familiar with the Swahili language and those with a history of neurological disorders were not eligible to participate in the study. Handedness of participants was recorded because in EEG research, it is often thought that being left-handed may have inherent neurophysiological differences than being right-handed. However, since the single electrode used in this experiment is close to the midline, the left-handed participants were included in the analyses.

Procedures

To begin, the participant was given a consent form to sign and was given the opportunity to ask any questions before the experiment began. Afterwards, the experimenter assisted the participant in putting the MindWave headset on and established an initial connection of the headset to the computer software. Upon successful connection, the experimenter let the system stabilize for approximately 90 seconds to allow for internal adaptation. After the connection was established, the participant began using the Java applet in which they were presented 17 Swahili words, one word at a time for ten seconds each word (for a total study duration of about three minutes). This 10 second study period for each item could not be shortened by pressing any keys, etc. This was to ensure that every study period was the same in duration. After completing the study period, the participant went directly into a testing period in which they were asked to type the English meaning (answer) of each presented Swahili word (e.g., a random Swahili word that was presented in the study period was presented alone in the test period, one at a time, to test recall). There was no time limit on how long they had to answer. Participants were allowed to type in any response including nonsense words or press the enter key to move on to the next item if they felt they did not know the answer. No feedback was given in the testing period.

Results

Data Analyses

Upon inspection of each individual's EEG frequency distribution, distributions tended to be skewed. All duplicate values (due to hardware lag) were removed by rejecting any individual one second record that contained the same values as the one second record before it. The problem with the MW was similar to video buffering; when the device could not process the EEG fast enough, it would repeat the previous recording. Extreme outliers were also removed by rejecting the EEG values when the sum of all eight frequency bands was higher than 3,000,000 (which was 2% of the total data). Furthermore, for the remaining data, 41% of the data was rejected due to poor signal quality (any poor signal value that wasn't equal to 0 which indicated a good signal according to NeuroSky—see Appendix C for further explanation). Each EEG for each trial underwent a logarithm transform to normalize the distribution. Values with poor signal quality were rejected Averages and standard deviations of the logarithms for each participant's entire session were calculated. Normalization was achieved by computing *z*-scores for every participant's individual trials (17 word-pair trials). Finally, the average *z*-score was computed for each participant's correct and incorrect trials for every frequency band.

Using the mean *z*-scores (correct and incorrect) from the theta, beta, and gamma bands, a preliminary paired-samples t-test was conducted to assess the difference in the means of when participants correctly recalled word-pairs vs. when they did not in the theta, beta, and gamma bands. In the high gamma range (41-49.75 Hz) there was a statistically significant difference in high gamma power (indicated by the mean of the maximum *z*-scores across trials) when participants answered incorrectly ($M = 1.1, SD = .24$) compared to when participants answered correctly ($M = .96, SD = .30$), $t(33) = 2.00$, $p = .05$ (two-tailed). This was contradictory to what previous research has shown, so the data was visually inspected again to make sure there were not any errors or extraneous noise signals.

From each 10 s trial period, 2.0 s epochs were removed from the beginning of each EEG trial's records, because upon visual inspection these epochs showed artifacts which were likely a result of eye movement or blinking caused by the changing or "flickering" from one word-pair to the next (at the end of each trial, the screen instantaneously moved to the next word-pair). Past research has shown that these "flickering" visual stimuli could cause resonance frequencies and spikes common to the gamma range, and removing these epochs is thought to be valid and common practice in

EEG data analyses (Hermann, 2001). After removal of these epochs, this cleaned data set was used for all further analyses.

Aggregate Within-Subjects Results

A paired-samples *t*-test was conducted to evaluate the difference in gamma, theta, and beta band frequency power when participants recalled word-pairs correctly and when they failed to recall word-pairs correctly. In the high gamma range (41-49.75 Hz) there was a statistically significant difference in gamma power (indicated by mean *z*-score of gamma maximum values) when participants answered incorrectly ($M = 1.02$, $SD = .24$) compared to when participants answered correctly ($M = .83$, $SD = .31$), $t(33) = 2.70$, $p = .01$ (two-tailed). The mean difference between "correct" mean z-scores of gamma maximums and "incorrect" mean *z*-scores of gamma maximums was .19 with a 95% confidence interval ranging from .05 to .33. The eta squared statistic (.18) indicated a large effect size. These results indicate that gamma power was higher when incorrectly remembered word-pairs were studied than when correctly remembered word-pairs were studied. In the low gamma range, there was no significant difference between the correct mean *z*-score of gamma maximums ($M = .87$, $SD = .41$) and incorrect mean *z*-score of gamma maximums ($M = .95$, $SD = .41$), $t(33) = 1.21$, $p = .23$ (two-tailed). A summary of the comparisons is shown in Table 1.

Table 1

Comparison of Correct Mean Z-Scores vs. Incorrect Mean Z-Scores

Frequency Band	**Correct Mean *z*-Score**	**Incorrect Mean *z*-Score**	***SD***	***t*-value**	***df***	***p*-value (2-tailed)**
Theta	-.04	.03	.30	1.29	33	.21
Low Beta	.02	.03	.32	.293	33	.77
High Beta	-.02	.03	.37	.643	33	.52
Low Gamma	.87	.95	.41	1.21	33	.23
High Gamma	**.83**	**1.02**	**.40**	**2.70****	**33**	**.01****

*Note. n = 34, **p=.01, two-tailed.*

In the theta range (3.5-6.75 Hz) there was no statistically significant difference in theta power (indicated by mean *z*-score) when participants answered incorrectly (*M = .03, SD = .22*) compared to when participants answered correctly (*M = -.04, SD = .21*), *t*(33) = 1.29, *p* = .20 (two-tailed). Our hypothesis that theta wave power would be significantly different for when correctly remembered word-pairs were studied vs. when incorrectly remembered word-pairs were studied was not supported by these results.

In the low beta range (13-16.75 Hz) there was no statistically significant difference in beta power (indicated by mean *z*-score) when participants answered incorrectly (*M = .03, SD = .19*) compared to when participants answered correctly (*M = .02, SD = .18*), *t*(33) = .293, *p* = .77 (two-tailed). In the high beta range (18-29.75 Hz) there was no statistically significant difference in power (indicated by mean *z*-score) when participants answered incorrectly (*M = .03, SD = .20*) compared to when participants answered correctly (*M = .02, SD = .25*), $t(33) = .643$, $p = .52$ (two-tailed). Again, our hypothesis that beta power would be significantly different for when correctly remembered word-pairs were studied vs. when incorrectly remembered word-pairs were studied was not supported by these results.

Exploratory Analyses

To further understanding of the NeuroSky MW's capability and possible usage, we conducted some exploratory analyses that helped identify possible hypotheses in future research using the device. First, we computed Spearman correlations in each frequency band for each user's *z*-scores and individual trial recall (see Appendix D). Of these tests, 13.6 is 5% (expected by chance) of the 272 tests computed. We found a total of 20 significant correlations (7 positive and 13 negative) of frequency band *z*-scores and correctly recalled word-pairs within nine different individuals (see Appendix D). Since multiple tests were run, we conducted an omnibus chi-square test to assess the significance of the difference of our observed vs. expected significant correlation count. The test indicated no significant difference, χ^2 (1, $n = 272$) = 3.01, $p < .10$ and $p > .05$. However, this seemed to indicate that there was a trend towards significance so we included more data to find if that trend continued. We conducted an additional omnibus chi-square test at the $p < .20$ level for the correlations in attempt to use more of the data to answer the same question, but the test again indicated no significant difference, χ^2 (1, $n = 272$) = .238, $p < .75$ and $p > .50$. The *p*-value moved away from significance suggesting that the trend we originally saw most likely occurred by chance due to the low *n* expectation. These results indicate that through our investigation of individual differences at the trial level, this particular device is probably not capable of generating usable data to construct an adaptive personal learner model for individuals at this time—at least not based on individual frequency bands.

A logistic regression was also conducted and can be found in Appendix F. The main goal of these exploratory analyses was to provide data and insight that will help drive future hypotheses in personal learner modeling using BCIs. It is stressed that these exploratory results should be considered with caution and that further experimentation needs to be conducted in order to assess the MW's (as well as other BCIs') potential in personal learner modeling.

Discussion

In the aggregate within-subjects analysis, it was discovered that there was a significant difference in the mean z-scores of gamma frequency power maximums (in the range of 41-49.75 Hz) from when participants studied correctly remembered word-pairs and when they studied incorrectly remembered word-pairs. However, the results indicated that the direction of the obtained effect was opposite of what previous research has found. Instead, these results suggest that when people remembered words incorrectly, their relative gamma power maximums were larger in the study period than in the study period when they remembered the words correctly. In the exploratory analyses, we found some weak but interesting results, and our investigation may lend insight to future research using BCIs.

Eye Movement

While experts do agree that gamma band activity is correlated to cognitive processing, they still do not agree on specific meaning of a gamma band response. Some EEG researchers are not convinced that gamma band responses are recorded accurately.

In fact, the most commonly reported gamma band response is thought by some to be a "…product of eye movements, probably of muscular origin, and not a direct measure of neuronal oscillations" (Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008). Yuval-Greenberg et al. used video based eye tracking and traditional EEG concurrently to investigate eye movement, and they concluded that when saccadic movement was included in the EEG, gamma spikes (or induced-gamma band responses) were consistently present as opposed to when saccadic movement was not included. Yuval-Greenberg et al. (2008) concluded the following:

> Most studies reporting induced Gamma Band Response (iGBR) present a stimulus following a fixation period. The appearance of a new stimulus (frequently with no explicit fixation point) starts the saccadic inhibition-enhancement sequence, involving mostly very small saccades of the microsaccade and intrusion type. Saccades are invariably accompanied by the spike potential, which, because of its short duration, translates into a relatively wide-band high-frequency activity in the spectral domain. (p.10)

So, considering eye movement, it seems plausible for there to be a higher gamma band power maximum on average for when people studied words they did not recall in a later test, because it may suggest that during those trials people exhibited more eye movement—they may have looked around the room more or may have not explicitly attended to the word-pairs as when they were perhaps more fixated (while exhibiting some but less saccadic behavior) on the word-pairs that they did remember.

This raises the question of whether or not these saccades are predictive of when someone is not attending to a stimulus or not. If these "spike potentials" always accompany saccadic behavior, then maximum gamma power values might be able to suggest when someone will not remember a word-pair. These notions would make good research questions for future research using the MW. These types of results might only be specific to the MW because of its electrode placement near the eyebrow, and may or may not replicate using other BCIs. One possible reason this study did not get the same type of results (increase in gamma for correctly recalled items) as previous research is the fact that these spikes might have contaminated the data set because of the inability of the MW to filter these spikes out. The MW does have a blink strength indicator which reports relative intensity of the most recent blink and ranges from 0 to 255. However, we did not find this to be a reliable indicator of eye movement or noise contribution due to its high degree of variability and arbitrary unit of measurement (has no units). Also, it seems that muscular activity such as head movement and other normal body movements contribute noise in the data.

Personal Learner Modeling in the Future

Based on our results, our idea of adaptive personal learner modeling might not yet be feasible using the MW. However, acquiring more participants might aid in the discovery of additional useful information from the MW in regards to personal learner modeling, and more research using the MW might help answer the question of whether personal learner models are achievable from its EEG data. Alternatively, another experiment which uses several paired-word sets could implement several study periods followed by test periods in order to track individual EEG over a longer period of time. If a personal signature of EEG activity does in fact exist in each user, then tracking their data over a long period of time (which includes replications) could help in determining the practicality of personal learner modeling using the MW. Research using other BCIs that have more spectral resolution (more electrodes) and provide more usable data may provide further information regarding possible personal EEG signatures.

MindWave

Considering the non-significant results regarding the theta and beta band hypotheses and the significant gamma results, it seems that in this study, we have uncovered some flaws with the MW. Its clear limitation is the fact that it only has one electrode. The placement of that electrode may not be the best since it is highly susceptible to muscle movements of the face, especially near the eyebrow. Another study could be conducted using the MW in which participants could be positioned to stay fixated on the stimulus. However, any results found would not be generalizable to any naturalistic situation—because in the classroom, students move normally. This makes the device impractical for use in any situation in which movement is not tightly controlled. We found that the MW had a tendency to lag, which resulted in poor resolution at times as many power values had to be filtered out for analyses. This is a clear cause for concern because for every missing value the device fails to give, it reduces the precision and accuracy of the EEG (41% of our total dataset was rejected due to poor signal quality).

In summary, the MW's reliability and accuracy remains unknown, and further testing is necessary for a better determination of what practical uses the MW is capable of supporting. Further research using the MW and other BCIs is definitely necessary in determining if these devices have the capability of indicating when students are learning. Based on this study's results, this particular device is not ready to move into the classroom, and more research needs to be conducted to determine its efficacy in pedagogy and in research.

General Limitations

If the MW is capable of accurate recording, another possible reason that similar results to previous research were not found could be the difference in task selection. This study's task only included one study followed by a simultaneous testing period. Future studies may benefit from including a retention period, and an additional review of the word-pairs. Since most previous research studies which found gamma relationships with encoding had longer training periods, it may be that we found a shorter-term gamma relationship that is specific to the current task and not seen in longer-term memory tasks. Maybe, gamma activity decreases upon initial encoding and increases later, possibly in a retention period. Additional experiments could test this speculation by recording and tracking an individual's gamma activity over an entire training period as opposed to an individual section of a task, e.g., a study period. Additional experiments that record EEGs in a testing period and analyze differences and similarities in study periods and testing periods might be useful as well. For example, if individuals show personal EEG signatures, comparing encoding periods to recall periods may provide critical insight into developing personal learner models.

As mentioned earlier, recording with one only electrode decreases the EEG's accuracy and precision, and it may not give enough information needed to identify specific cognitive processing. Some BCIs are commonly known to have a poor signal to noise ratio, and they are considered non-stationary because EEG signals vary over time and over sessions (Lotte, Congedo, Lécuyer, Lamarche, & Arnaldi, 2007). The main limitation seems to be the MW's partiality to movements in general. Head nodding, head turning, coughs, sneezes, eye blinks, etc. all produce noise in the EEG, and this has been shown in repeated testing that compared when people blinked and acted normally, and when people held their eyes open and tried to be as still as possible for a recording session (Larsen, 2011). Larsen (2011) proposed that blinks could be detected more accurately by using a paradigm which factors in delta, theta, alpha, and gamma band activity. Larsen's study (2011) investigated the classification of EEG signals using the NeuroSky Mindset (has the same chipset as the MW), and suggested that blinks occurred when a simultaneously high boost of delta, theta, and alpha amplitudes occurred along with a decrease in gamma amplitude. Taking these suggestions into consideration, future advancement of this particular device could be done to develop a better filtering technique which might help filter out noise caused by movements. If such filtering could be done, this experiment could be replicated to assess the validity of its results.

Conclusion

Although the MW is inexpensive and has been touted by NeuroSky as being a great research tool for detecting attention and meditation, this study suggests that the MW is not quite advanced enough to give reliable and accurate scientific research data for specific research purposes (e.g., identifying certain cognitive processes) based on its inability to corroborate prior, replicated results from EEG literature. A follow-up experiment should be conducted and could replicate this study (using the same PAL task and same word-pairs) but use a different BCI to determine the validity of this study's results. Several BCI options exist, such as the esteemed Emotiv Epoc which features 14 electrodes, but many require saline-gel and cost more than the MW. As the world advances technologically, BCIs are sure to improve significantly in design, accuracy, and precision, and could possibly stake their claim as an important research tool in the scientific community, but only after further testing and research has been carried out. It could be possible that this experiment has uncovered some novel phenomenon in which a brief decrease occurs in the high gamma range in a very short time period, but several additional studies would need to replicate these findings.

Additionally, our intent was to push for a move into the classroom to help student learning. Although this particular device does not seem to be robust enough to be used for determining if students are learning or not as of now, it does provide some interesting data. If higher *z*-scores of the high gamma maximum values are indicative of recall failure using the MW, then this raises many speculations. The most interesting is that the device seems to be useful in capturing muscle movements, such as excessive eye movement. Several physiological sensors try to lend evidence of when certain behaviors are taking place, and the MindWave seems to possess an uncanny ability to record saccadic behavior quite well. In this sense, the MW is a very useful tool in determining when people are not focused based on the muscle movement spikes it has a tendency to record as a result of being right above the eyebrow. Just as Viagra was originally used for treating pulmonary arterial hypertension and was fortuitously found to also treat erectile dysfunction, it could be that we have discovered another use for the MW. Could it be used to detect muscle movements? Are muscle movements what are mainly recorded by the MW? These questions could be tested by replicating this study adding video-based eye tracking to compare the differences in saccadic movement and non-saccadic movement groups, e.g., one group fixates on word-pairs explicitly while another group looks around the room explicitly. Future work with portable EEG headsets (especially including the MW) should be conducted to uncover their potential as useful tools in today's fast-growing technological world.

References

Asaad, W. F., Rainer, G., & Miller, E. K. (1998). Neural activity in the primate prefrontal cortex during associative learning. *Neuron, 21*, 1399-1407.

Bardouille, T., Picton, T. W., & Ross, B. (2010). Attention modulates beta oscillations during prolonged tactile stimulation. *European Journal of Neuroscience, 31*, 761-769.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, *87*, 527-570.

Bouyer , J. J., Montarom, M. F., Vahnee, J. M., Albert, M. P., & Rougeul, A. (1987). Anatomical localization of cortical beta rhythms in cat. *Neuroscience*, *22*, 863-869.

Chang, K. M., Nelson, J., Pant, U., & Mostow, J. (2013). Toward exploiting EEG input in a reading tutor. *International Journal of Artificial Intelligence in Education, 22*, 19-38.

Choi, H., Jones, A., & Schwartz, G. (2012). Using brain-computer interfaces to analyze EEG data for safety improvement. Retrieved from Team for Research in Ubiquitous Secure Technology website: http://www.truststc.org/education/reu/12/Papers/Choi_Jones_Paper.pdf.

Cox, D. R., & Snell, E. J. (1989). Analysis of binary data. CRC Press.

Crowley, K., Sliney, A., Pitt, I., & Murphy, D. (2010, July). Evaluating a brain-computer interface to categorise human emotional response. In *Advanced Learning Technologies (ICALT), 2010 IEEE 10th International Conference on* (pp. 276-278).

DesJardins, S. (2001). A comment on interpreting odds-ratios when logistic regression coefficients are negative. *AIR Professional File*, *81*, 1-7.

Egner, T., & Gruzelier, J. H. (2004). EEG Biofeedback of low beta band components: frequency-specific effects on variables of attention and event-related brain potentials. *Clinical Neurophysiology*, *115*, 131-139.

Fu, K. M. G., Foxe, J. J., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2001). Attention-dependent suppression of distracter visual input can be cross-modally cued as indexed by anticipatory parieto–occipital alpha-band oscillations. Cognitive Brain Research, *12*, 145-152.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*, 374-385.

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., & Tutoring Research Group, University of Memphis (2000). AutoTutor: A simulation of a human tutor. *Cognitive Systems Research*, *1*, 35-51.

Gruber, T., Keil, A., & Müller, M. M. (2001). Modulation of induced gamma band responses and phase synchrony in a paired associate learning task in the human EEG. *Neuroscience letters*, *316*, 29-32.

Gruber, T., Tsivilis, D., Montaldi, D., & Muller, M. M. (2004). Induced gamma band responses: an early marker of memory encoding and retrieval. *Neuroreport*, *15*, 1837-1841.

Hebb, D. O. (2002). The organization of behavior: *A neuropsychological theory*. Psychology Press.

Herrmann, C. S. (2001). Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. *Experimental brain research*, *137*, 346-353.

Herrmann, C. S., Arnold, T., Visbeck, A., Hundemer, H. P., & Hopf, H. C. (2001). Adaptive frequency decomposition of EEG with subsequent expert system analysis. *Computers in Biology and Medicine, 31(6),* 407–427.

Herrmann, C. S., Fründ, I., & Lenz, D. (2010). Human gamma-band activity: A review on cognitive and behavioral correlates and network models. *Neuroscience and Biobehavioral Reviews*, *34*, 981-992.

Herrmann, C. S., Lenz, D., Junge, S., Busch, N. A., & Maess, B. (2004). Memory-matches evoke human gamma-responses. *BMC Neuroscience, 5,* 5-13.

Herrmann, C. S., Mecklinger, A., & Pfeifer, E. (1999). Gamma responses and ERPs in a visual classification task. *Clinical Neurophysiology*, *110 (4),* 636–642.

Hoogenboom, N., Schoffelen, J., Oostenveld, R., Parkes, L. M., & Fries, P. (2006). Localizing human visual gamma-band activity in frequency, time and space. *Neuroimage*, *29*(3), 764-773.

Jensen O. & Tesche C. D. (2002) Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience, 15*, 1395–1399.

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9 - 12 Hz) in - crease with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*(8), 877-882.

Klimesch, W., Doppelmayr, M., Yonelinas, A., Kroll, N. E., Lazzara, M., Röhm, D., & Gruber, W. (2001). Theta synchronization during episodic retrieval: Neural correlates of conscious awareness. *Cognitive Brain Research*, *12*(1), 33-38.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology*, 45, 812-863.

Larsen, E. A. (2011). *Classification of EEG Signals in a Brain-Computer Interface System* (Doctoral dissertation, Norwegian University of Science and Technology).

Lotte, F., Congedo, M., Lécuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain–computer interfaces. *Journal of Neural Engineering, 4.*

Marshall, L., Helgadóttir, H., Mölle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature, 444*, 610-613.

Miltner, W. H., Braun, C., Arnold, M., Witte, H., & Taub, E. (1999). Coherence of gamma-band EEG activity as a basis for associative learning. *Nature, 397*, 434-436.

Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*, 691-692.

Niedermeyer, E., & Da Silva, F. L. (2005). *Electroencephalography: Basic principles, clinical applications, and related fields*. Philadelphia, PA: Lippincott Williams & Wilkins.

O'Nuallain, S. (2009). Zero power and selflessness: what meditation and conscious perception have in common. *Cognitive Sciences 4*, 49-64.

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *The Journal of neuroscience, 26*, 7523-7531.

Pantev, C. (1995). Evoked and induced gamma-band activity of the human cortex. *Brain Topography*, 7*,* 321-330.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology*, *49*, 1373-1379.

Pulvermüller, F., Lutzenberger, W., Preil, H., & Birbaumer, N. (1995). Spectral responses in the gamma-band: physiological signs of higher cognitive processes? *Neuroreport*, *6*, 2059-2064.

Rani, P., Sarkar, N., & Liu, C. (2005). Maintaining optimal challenge in computer games through real-time physiological feedback. In *Proceedings of the 11th International Conference on Human Computer Interaction* (pp. 184-192).

Rebolledo-Mendez, G., Dunwell, I., Martínez-Mirón, E. A., Vargas-Cerdán, M. D., De Freitas, S., Liarokapis, F., & García-Gaona, A. R. (2009). Assessing NeuroSky's usability to detect attention levels in an assessment exercise. In *Human-Computer Interaction. New Trends* (pp. 149-158). Springer Berlin Heidelberg.

Rebolledo-Mendez, G., & De Freitas, S. (2008). Attention modeling using inputs from a brain computer interface and user-generated data in Second Life. In *The Tenth International Conference on Multimodal Interfaces (ICMI 2008), Crete, Greece*.

Sederberg, P.B., Kahana, M.J., Howard, M.W., Donner, E.J., & Madsen, J.R., (2003). Theta and gamma oscillations during encoding predict subsequent recall. *Neuroscience*, *23,* 10809–10814.

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science*, *153*, 652-654.

Sternman, M. B., & Egner, T. (2006). Foundation and practice of neurofeedback for the treatment of epilepsy. *Applied Psychophysiology and Biofeedback*, *31*, 21-22.

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, *2*, 1-11.

Tiitinen, H. T., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40 Hz transient response in humans. *Nature*, *364*, 59-60.

Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., & Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron, 58(3)*, 429-441.

Appendix A

Swahili-English Word-Pairs Used

Swahili Word	***English Meaning***
Asante	Thank you
chakula	food
Habari	Good Morning
samaki	fish
ukuta	wall
tano	five
leo	today
manjano	yellow
Kwaheri	Goodbye
Mimi	I
tembo	elephant
mamba	alligator
kuku	chicken
bafu	bathroom
Babu	Grandfather
viatu	shoes
rafiki	friend

Appendix B

Comparison of Mean Z-Scores for All Frequency Bands

Frequency Band	**Correct Mean *z*-Score**	**Incorrect Mean *z*-Score**	***SD***	***t*-value**	***df***	***p*-value (2-tailed)**
Delta	-.04	.05	.35	1.47	33	.15
Theta†	-.04	.03	.30	1.29	33	.21
Low Alpha	-.01	.03	.29	.657	33	.52
High Alpha	-.02	.05	.30	1.43	33	.16
Low Beta†	.02	.03	.32	.293	33	.77
High Beta†	-.02	.03	.37	.643	33	.52
Low Gamma†	.87	.95	.41	1.21	33	.23
High Gamma†	.83	1.02	.40	2.70**	33	.01**

*Notes. n = 34, **p=.01, two-tailed.* † hypothesized bands.

Appendix C

As found in the NeuroSky User Manual, below is a brief description of the MindWave's output measures.

Description of NeuroSky MindWave Frequency Band Ranges and Output Measures

Delta	0.5 - 2.75Hz
Theta	3.5 - 6.75Hz
Low Alpha	7.5 - 9.25Hz
High Alpha	10 - 11.75Hz
Low Beta	13 - 16.75Hz
High Beta	18 - 29.75Hz
Low Gamma	31 - 39.75Hz
High Gamma	41 - 49.75Hz
Meditation (eSense meter)	Returns the eSense Meditation integer value, between 0 and 100
Attention (eSense meter)	Returns the eSense Attention integer value, between 0 and 100
Poor Signal	Returns poor signal level, 0 is good signal, 200 is off-head state
Blink Strength	Returns an integer value between 0-255, indicating the blink strength

Appendix D

Spearman Correlations and *p*-values

User	Delta	High Alpha	Low Beta	High Beta	Low Gamma	High Gamma
User 1	-.600, *p*=.014		-.512, *p*=.043	-.541, *p*=.03		
User 6				-.529, *p*=.029		
User 14	-.512, *p*=.035	-.488, *p*=.047		-.537, *p*=.026	-.488, *p*=.047	-.634, *p*=.006
User 18				-.527, *p*=.03		
User 19			.537, *p*=.026	.561, *p*=.019		
User 24	-.729, *p*=.001					
User 29	.679, *p*=.003	.623, *p*=.008	.538, *p*=.026	.679, *p*=.003	.538, *p*=.026	
User30	-.504, *p*=.039					
User 31				-.708, *p*=.001		

Note. n=272

Appendix E

Logistic Regression

As a final exploratory analysis, a forward stepwise logistic regression was performed to assess the impact of several factors within every frequency band (shown in Table 2) on correctness of recall of the Swahili word-pairs.

Table 2

Description of Subject and Trial Level Variables used in Logistic Regression

Trial Level	**Subject Level**
Log averages of each frequency band for each individual trial for each user	Averages of each frequency band for all trials of each user
Log average of low gamma to log average of theta ratio for each individual trial for each user	*SD*s of each frequency band for all trials of each user
	Gender
Log average of low beta to log average of delta ratio for each individual trial for each user	Age
	Handedness
Log average of high alpha to log average of beta ratio for each individual trial for each user	
Z-Scores of every frequency band for each individual trial for each user	

The forward stepwise regression's best-fitting model contained eight independent variables (the standard deviations of delta, low beta, high beta, and high gamma, the high gamma *z*-scores, the high alpha *z*-scores, the mean high alpha, and the mean low gamma).

The found model containing the abovementioned eight predictors was statistically significant, χ^2 (8, N = 585) = 62, $p < .001$. Since, this regression did not meet the proposed criterion for stable results, which suggests having at least 10 participants per predictor (Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996), the model needed cross-validation. The model was validated five times using ten-fold cross-validation in which 10% of the sample was randomly selected and tested using the found model ten times. The results are shown below in Table 3.

Table 3

Cross Validation Results

Model	**Train**	**Test**
	Spearman *r (SE)*	Spearman *r (SE)*
BestFit	.33991 (.00021)	.24204 (.00196)

Note. n = 50

The model as a whole explained between 12% (Cox & Snell, 1989) and 16% (Nagelkerke, 1991) of the total variance in correct recall of word-pairs and correctly classified 64.3% of cases. Although these percentages are low, we found them to be reliable and statistically significant (*see Table 4 below*). Of the eight variables included in the model, seven made a statistically significant contribution to the model (all except the standard deviation of delta were significant—however, the standard deviation of delta had a p-value that was very close to significance, $p = .08$).

Table 4

Forward Stepwise Logistic Regression Predicting Correctness of Recall

Variable in Model	**Beta**	**Wald**	***df***	***p***
SD low beta††	1.08	2.69	1	.006**
Mean high alpha††	-3.84	-6.19	1	<.001***
SD high gamma††	-2.06	-5.56	1	<.001***
Z-score high alpha†	2.97	5.32	1	<.001***
SD high beta††	1.76	3.62	1	<.001***
SD delta††	.50	1.74	1	.08
Z-score high gamma†	-.67	-2.82	1	.004**
Mean low gamma††	.85	2.55	1	.01**
Constant		4.49		<.001***

Notes. Standardized Coefficients. Refer to Table 2 for complete description of variables. † Trial level variable. †† Subject level variable. **$p \leq .01$. ***$p < .001$.

Interpreting the standardized coefficients (see Appendix E for unstandardized coefficients), the strongest predictor of correct recall was the mean of high alpha, a subject level variable, β = -3.84. This result suggests that as a participant's average high alpha power for the entire experiment (17 trials) changed by one standard deviation, their recall of word-pairs would change by -3.84 standard deviations. If this log of odds coefficient is converted to an odds ratio by exponentiating the inverse can be interpreted most easily. This suggests that when the mean of high alpha is one standard deviation higher, there is a 98% decrease in odds of recalling correctly.

Interestingly, the second strongest predictor(s) were the *z*-scores of high alpha, which are trial level variables, $\beta = 2.97$. This suggests that as a participant's *z*-score of high alpha for an individual trial changed by one standard deviation, their recall of word-pairs would change by 2.97 standard deviations. These two predictors suggest that an overall increase in high alpha power for the user during the entire experiment (all 17 trials) is predictive of lower recall, while an increase in a user's high alpha *z*-score in a single trial is predictive of higher recall. These peculiar results could be thought of like so: An overall sustained, average high alpha power is not good for recall, but a brief increase during a single trial is good for recall. This might be indicative of visual system taxation (Fu et al., 2001). If a user's visual system is taxed too greatly, it might interfere with encoding. There could be an optimal range for high alpha band power in regard to encoding new information, and future research could investigate this.