Empirical Evaluation of the Emotiv EPOC BCI Headset for the Detection of Mental Actions

Grant S. Taylor and Christina Schmidt
University of Central Florida, Institute for Simulation & Training

This study evaluated the detection accuracy of one of the first brain-computer interfaces intended for personal use by normal, healthy users: the Emotiv EPOC. This system allows the user to directly interact with computer software through thoughts alone. The system was evaluated on its ability to accurately detect and classify six sets of paired mental actions over three evaluation phases. Results found the system to perform significantly better than chance for all mental actions, and improve over time with additional training data. The system detected all actions with equivalent accuracy. Additional investigation of individual differences revealed that a user's gender, age, handedness, attentional control, vividness of visual imagery, and mental rotation ability all had no bearing on the detection accuracy of the system. These results indicate the Emotiv EPOC system performs its function as a brain-computer interface with an acceptable level of accuracy, yielding many new possibilities for human-computer interaction.

 DOI 10.1177/1071181312561017

INTRODUCTION

Background of Brain-Computer Interfaces

Brain-computer interfaces (BCIs) utilize psycho-physiological measurement techniques to record brain activity and translate it into actions through a computer system. This allows the user to interact with a system through mental actions alone, circumventing traditional control paradigms such as physical manipulation or verbal commands (Kubler, Kotchoubey, Kaiser, Wolpaw, & Birbaumer, 2001). The methods used to monitor the user's brain activity can be either noninvasive (such as electroencephalogram (EEG)) or invasive (cortically-implanted electrodes). Invasive techniques provide more precise and detailed measures, but come at the cost of requiring major surgery.

Regardless of the method used, the stream of data recorded from the user must be processed by an algorithm designed to decode the user's intentions. This algorithm can be relatively robust and generalizable across users if the system relies on simple event-related potentials (ERPs). However, this limits the system's functionality. More advanced signal processing algorithms are required to yield more complex system capabilities.

The implementation of BCIs has traditionally been limited to clinical use, providing "locked in" patients suffering from physical trauma to the central nervous system or diseases such as Amyotrophic Lateral Sclerosis, Cerebral Palsy, or Parkinson's with a means to interact with the world (Daly & Wolpaw, 2008). These patients are incapable of the motor functioning necessary to control traditional computer input devices, or even speak, and so BCIs are typically their one and only option to interact with the world and communicate with others. In this way, BCIs provide an invaluable service for locked in patients, allowing them to regain a modest degree of normal human interaction that would otherwise be impossible. For this reason, clinical patients are willing to tolerate the shortcomings of BCIs: relative to traditional control schemes BCIs are slow, limited in their functionality, and sometimes highly invasive.

However, continued advances in both sensor technologies and signal processing techniques are reducing these shortcomings, questioning whether BCIs should still remain solely as assistive devices in the clinical realm, or whether they are ready to become a viable option for a variety of applications by the general population.

The Emotiv EPOC BCI System

The Emotiv EPOC headset is one of the first commercially-available BCI systems intended for normal, healthy users. While traditional BCI systems cost several tens of thousands of dollars, the EPOC is currently priced at only $300 for a consumer model (developer and researcher versions are available at greater cost), making it much more affordable for any potential user. It also has the benefit of requiring virtually no technical expertise, with system setup and the computer software designed for novice users.

The EPOC headset consists of 16 sensors that record electrical activity originating from the brain at the scalp. The resulting data is transmitted wirelessly to a receiver unit connected to a standard computer. Here, the signal is processed using one of Emotiv's three detection suites: *Expressiv* detects the user's facial expressions, *Affectiv* detects emotional state, and *Cognitiv* detects conscious thoughts and intentions. Once detected, the user can program the software to trigger virtual keytrokes. For example, the system can be set to press the "A" key when the user thinks "left" and the "D" key when they think "right", allowing them to mentally control a simple video game.

With these capabilities at such minimal cost, it seems that the EPOC headset could bring a whole new method of computer interaction to regular users. However, its capabilities must first be thoroughly evaluated. Given the system's dramatically reduced price relative to similar research-grade EEG systems, users initially questioned the system's ability to accurately record brain activity. Although this concern seems to have been alleviated (Ekanayake, 2011), the system's ability to accurately detect and classify user expressions, emotions, and thoughts must still be investigated.

Preliminary evaluations of the EPOC headset have consisted primarily of proofs-of-concept rather than true empirical research. For example, Ranky and Adamovich (2010) tested the EPOC's ability to control an external robotic arm. Although they found that users were capable of successfully controlling the robotic arm, no data regarding its ability to accurately detect mental actions was recorded.

Lievesley, Wozencroft, and Ewins (2011) conducted a controlled system evaluation of the EPOC's ability to detect facial expressions and conscious thoughts. However, this study is only capable of providing preliminary evidence given its miniscule sample size (N = 3). Additional research is necessary to empirically evaluate the EPOC's ability to perform as a BCI device.

This was the aim for the present experiment. Specifically, this empirical research was conducted independent of any EEG manufacturer as a confirmatory study for assessing the utility of a low-cost, low-density BCI system for future application for the general public and Warfighter systems.

Experimental Overview

This experiment evaluated the EPOC's *Cognitiv* detection suite, which is intended to detect a user's mental commands (e.g. push, pull, rotate left, rotate right, etc.). Analyses determined the overall accuracy of the system, the relative accuracy of each individual command, the benefit of repeated training phases, and the impact of individual differences (such as gender and handedness) on system accuracy. This information will provide future users with a more thorough understanding of the system's capabilities and limitations, allowing for the device to be used more effectively in research, system design, and personal use.

METHODS

Participants

The participants in the study were 57 (34 male, 23 female) university undergraduate students (age: M = 19.96, SD = 4.34). Each participant was compensated with course credit.

Materials

Emotiv EPOC headset. The EPOC is a consumer grade EEG. The outer shell is made of plastic and contains 16 sensors that record electrical signals from the scalp. Two of these sensors (placed directly above each ear) serve as references. Each sensor is saturated with a sterile saline solution to enhance conductivity. The signal from the headset is transmitted wirelessly to a standard computer, where it is processed by the Emotiv software.

Emotiv software. The software, which records and processes the data received by the headset is developed specifically for the EPOC headset by Emotiv. This software automatically generates detection algorithms for each mental action (Table 1) based on training data provided by the user. Once the algorithms are developed, the software monitors the incoming signal for the presence of one of the trained actions. When an action is detected, the software triggers a user-defined virtual keystroke (equivalent to physically pressing a key on a keyboard), to be interpreted by other software running on the computer. The EPOC system is available in three forms intended for consumers, developers, and researchers. All systems provide the same headset and basic detection software, with the developer version including a software development kit (SDK), and the researcher version allowing access to the raw EEG data stream. This evaluation used only the consumer version of the system, utilizing the detection algorithms that are consistent among all versions.

Push	Pull
Lift	Drop
Left	Right
Rotate clockwise	Rotate counter-clockwise
Rotate forward	Rotate reverse
Rotate left	Rotate right

Table 1. Pairs of mental commands available in the EPOC's Cognitiv detection suite. One additional action is available (Disappear), but was excluded from the study because it lacked a paired action.

System evaluation software. A computer program was developed specifically for this experiment to present instructions to participants and record the detection output from the Emotiv software. This program instructed participants which mental action to focus and provided feedback regarding the detections made by the Emotiv software (Figure 1).

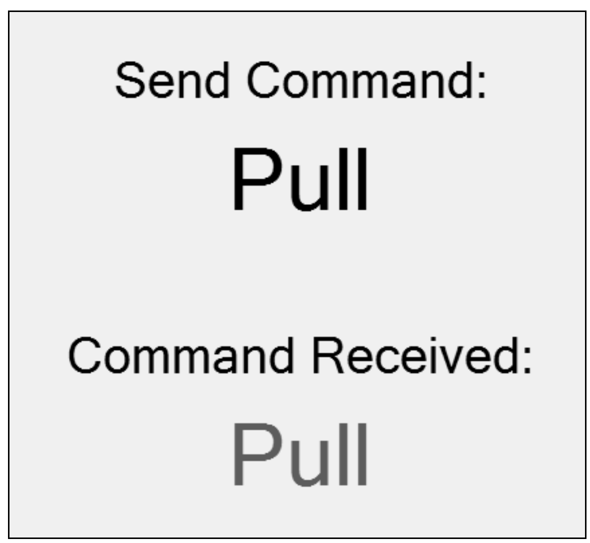

Figure 1. Sample trial from the evaluation software.

Additional measures. Participants completed several additional measures considered to be potentially relevant to their ability to successfully utilize the EPOC system.

A demographics questionnaire recorded participant age and gender, as both are known to have an impact on the physiological structure of the brain (Hsu et al., 2008).

A self-report handedness measure (McManus, 2009) was used to classify participants along a continuum between purely right-hand dominant and purely left-hand dominant, given the known impact of handedness on brain structure (Morgan & McManus, 1988).

A self-report measure of attentional control (Derryberry & Reed, 2002) was used to record participants' ability to focus attention, based on the prediction that those with greater attentional control may provide a more consistent pattern of neural activity, which would improve the system's ability to detect mental actions.

The vividness of visual imagery questionnaire (Marks, 1995) allowed participants to rate their ability to imagine visual scenes. Given that the EPOC system relies on the participant imagining the movement of an item, this metric was recorded with the expectation that those with greater visual imagery vividness would achieve better accuracy with the system.

The cube comparison test (Vandenberg & Kuse, 1978) was employed to objectively measure participants' mental rotation ability. This was tested given the prevalence of rotation-based mental commands used by the system, which may provide an advantage to those with greater mental rotation abilities.

Procedure

After the participant read and agreed to an informed consent form, they completed the demographics, handedness, attentional control, and vividness of visual imagery questionnaires, followed by the cube comparison test. After the completion of these preliminary measures, the researcher placed the EPOC headset on the participant, following the procedure provided in the user manual. All sensors were confirmed to have acceptable impedance values before continuing. The participant was then provided with instructions of how to use the system. The researcher stressed the importance of maintaining focus on the mental action consistently throughout the training phase, as well as keeping their head and face muscles relaxed throughout the study to avoid having their muscle activity create artifacts in the EEG signal. This is important for any EEG system given the greater amplitude of the electrical signals from muscle activity relative to brain activity.

The participant then began the first of three training phases. Each phase began with the collection of 8 seconds of neutral brain data, during which the participant relaxed with their eyes open and not focusing on anything in particular. The participant then received instructions of how to train the first pair of mental actions. The mental actions were always grouped in the pairs listed in Table 1. The order in which the pairs of mental actions were trained was randomized across participants, as well as across training phases for single participants. During the first training phase, the researcher demonstrated the action the participant should imagine by moving a box in front of them. For example, for "push", the researcher slowly moved the box away from the participant, and for "pull" the researcher slowly brought the box toward the participant. Following these demonstrations, 8 seconds of data was collected for each action as the participant mentally focused on the action, allowing the software to generate its detection algorithms.

After the first pair of actions was trained, they were evaluated using the system evaluation software. The software instructed the participant when to focus on each of the two actions, as well as neutral periods when they should not focus on either. The software also provided the participant with real-time feedback regarding which action the system had most recently detected. For each evaluation phase, the participant completed a total of 18 randomly presented trials: 6 for each of the two mental actions and 6 neutral trials. Each trial lasted for 4 seconds with a 2 second inter-stimulus interval between each trial. During each trial, the evaluation software recorded every time an action was detected by the Emotiv software by recording the triggered virtual keystrokes. These keystrokes were programmed to be triggered continuously for the duration of the detected mental action, allowing the evaluation software to not only record that an action was detected, but also the duration of its detection. It is important to note that the Emotiv software was only monitoring for the presence of the two mental actions being evaluated in a given phase, limiting the classification to only three possibilities: either of the two actions or neutral.

Following the completion of the evaluation of the first pair of actions, the same procedure repeated for the training and evaluation of each of the remaining 5 pairs of actions to complete the first training phase. After a short break, the same procedure was repeated for training phases 2 and 3. For the second and third training phases, the Emotiv software does not generate entirely new detection algorithms based on the new training data, but rather combines the new data with old to refine the previously generated algorithm. After completing the third training phase the headset was removed from the participant and they were allowed to leave. The entire study lasted roughly 2 hours.

RESULTS

The system was evaluated by calculating two separate performance metrics: errors and false alarms. Errors occurred when the system incorrectly categorized a signal (e.g. a "push" error occurred if the user thought "push" but the system detected "pull"). False alarms occurred when the system detected a signal when none was intended (e.g. a "push" false alarm occurred when the user did not send an action, but the system detected "push"). Ratios were calculated from both values by dividing the total number of correct detections (e.g. the system detected "push" when the user thought "push") by the total number of either errors or false alarms for a given evaluation phase. Random chance performance would result in a value of 1 for both metrics (errors or false alarms occurred as often as correct detections), with larger values indicating better system performance. Using these metrics allows the system's classification accuracy (ability to discriminate one signal from another) and sensitivity (ability to discriminate signal from non-signal) to be evaluated separately.

Overall System Performance

Overall system performance was evaluated by calculating average values of error and false alarm ratios across all 6 pairs of actions and 3 training phases. Two-tailed one-sample *t*-tests were used to compare both values to a reference value of 1 to determine whether the system performed significantly better than chance. Both the error ratio [$t(56) = 10.534$, $p < .001$, $M = 7.02$, $SD = 4.31$] and false alarm ratio [$t(56) = 10.522$, $p < .001$, $M = 2.22$, $SD = 0.881$] were found to be greater than 1, indicating that the system's detection accuracy and sensitivity performed significantly greater than chance.

Action and Training Phase Comparison

The system's ability to detect each of the mental actions across each of the training phases was evaluated through a 12 (mental actions) * 3 (training phases) repeated measures ANOVA (Figure 2). No significant main effect was found for type of mental action for either error ratio [$F(11,616) = 1.407$, $p = .165$] or false alarm ratio [$F(11,616) = 1.446$, $p = .148$], indicating that the system was able to detect all actions with equivalent accuracy and sensitivity.

Training phase was found to have a significant main effect on false alarm ratio [$F(2,112) = 3.554$, $p = .032$], but its effect on error ratio was not significant [$F(2,112) = 0.660$, $p = .519$]. Post-hoc comparisons revealed that the false alarm ratio in the third training phase ($M = 2.619$, $SD = 1.895$) was significantly improved over the false alarm ratio from the first phase ($M = 1.944$, $SD = 0.966$, $p = .017$), while the second phase did not differ significantly from either phase one or three ($p > .05$ in each case).

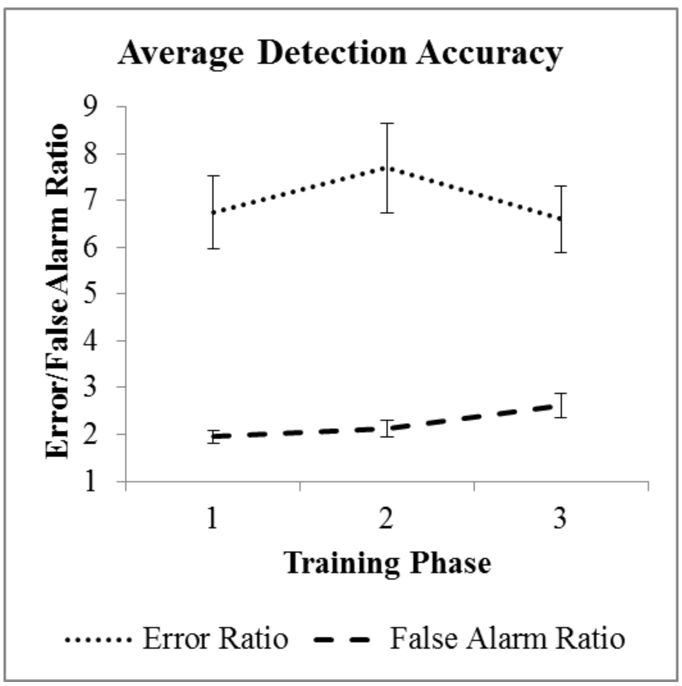

Figure 2. Average system detection accuracy for all mental actions. Higher values indicate better system performance for both metrics.

Individual Differences

Independent sample *t*-tests were conducted to compare males to females on globally averaged error and false alarm ratios. No significant differences were found ($p > .05$ in each case).

Correlations were computed to determine the relationship between system performance and participant age, handedness, mental rotation ability, attentional control, and vividness of visual imagery. No significant correlations were found ($p > .05$ in each case).

DISCUSSION

System Performance

This study sought to provide an empirical evaluation of the Emotiv EPOC's ability to detect and accurately classify mental actions, and determine whether various individual differences impact that ability. The results reveal that the system performed its function significantly better than chance. The system's sensitivity was fairly good, with correct detections occurring more than twice as often as false alarms, or an average of 45 false alarms for every 100 correct detections. Classification accuracy was more successful with correct detections occurring more than seven times as often as erroneous classifications, meaning the system was able to accurately classify mental actions 87.5% of the time. Although still far from the accuracy (or response time) that can be obtained through the use of physical interfaces such as keyboards or touchscreens, this performance data suggests that the system has reached a reasonable level of accuracy to be used for personal or research purposes.

Further evaluation revealed no significant differences in the system's ability to detect the 12 available mental actions. Therefore, users should expect equivalent system performance regardless of the mental actions they choose to use.

A significant effect was found for training phase, with the system's sensitivity improving by 35% from the first to the third phase. However, with the current study design it is impossible to differentiate whether this improvement is the result of the system refining its detection algorithms, or the user learning to provide more consistent patterns of activity with practice (or, more likely, some combination of both). More surprising is the fact that the system's detection accuracy did not improve similarly over time, indicating that the system reached a stable level of detection accuracy after only a single training phase.

Individual Differences

Surprisingly, none of the investigated individual differences were found to have a significant impact on system performance. This suggests that the system is robust against variation in neurophysiological structure, and that users' ability to focus, mentally rotate objects, or generate imagined visual imagery has no impact on their ability to successfully use the system. Although individual performance did vary considerably between participants, none of the seemingly obvious sources were responsible for this variation.

Limitations

The primary limitation of this study is the limited number of mental actions that were active in the detection software at any given time. During each evaluation phase, the system was only monitoring for the presence of two mental actions, and these two actions were always complete opposites (e.g. push and pull, or lift and drop). The system is capable of monitoring for additional actions simultaneously, and the inclusion of additional possible actions would almost certainly reduce the system’s detection accuracy.

Another limitation is the fact that participants only completed three training phases. Some previous research has included much more thorough training time. Additional training time would presumably allow for the system to further refine its detection algorithms while also allowing users to improve their ability to generate consistent signals, thus improving overall system accuracy.

One limitation of the EPOC system itself is the proprietary nature of the system and its software. The precise nature of the signal processing and detection algorithms is unknown to the user, which is understandable from a commercialization standpoint. However, this fact limits the system’s utility as certain high-stakes operational applications (such as the military domain) are unlikely to adopt a system based on anything less than complete transparent operation. The research domains capable of utilizing the EPOC system are also limited, as the lack of precise data processing details make the system of little benefit for neuroscientists primarily interested in the underlying neurological functions that the detection algorithms focus on.

Conclusions

Despite its reduced cost relative to traditional research-grade EEG systems, the EPOC system appears to be capable of recording brain activity for the detection of various mental actions successfully. This system will provide BCI capabilities to a whole new field of personal users, who use a BCI system not because it is a last and only option, but because it provides a novel and enjoyable means of interacting with technology. With continued system advancement and adoption by home users and researchers, this technology has the potential to bring about an entirely new field of human-computer interaction and usability considerations.

ACKNOWLEDGMENT

This research was conducted independently of Emotiv Systems. This work was supported by the US Army RDECOM (W91CRB08D0015). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of RDECOM or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

Daly, J. J. & Wolpaw, J. R. (2008). Brain–computer interfaces in neurological rehabilitation. *Lancet Neurology*, *7*, 1032-1043.

Derryberry, D. & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, *111*, 225-236.

Ekanayake, H. (2011). *P300 and Emotiv EPOC: Does Emotiv EPOC capture real EEG?* Retrieved from: http://neurofeedback.visaduma. info/emotivresearch.htm

Hsu, J., Leemans, A., Bai, C., Lee, C., Tsai, Y., Chiu, H., & Chen, W. (2008). Gender differences and age-related white matter changes of the human brain: A diffusion tensor imaging study. *NeuroImage*, *39*(2), 566-577.

Kubler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., & Birbaumer, N. (2001). Brain-computer communication: Unlocking the locked in. *Psychological Bulletin*, *127*(3), 358-375.

Lievesley, R., Wozencroft, M., & Ewins, D. (2011). The Emotiv EPOC neuroheadset: An inexpensive method of controlling assistive technologies using facial expressions and thoughts? *Journal of Assistive Technologies*, *5*(2), 67-82.

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery*, *19*, 153-167.

McManus, I. C. (2009). *The brief handedness questionnaire*. Retrieved from: http://www.ucl.ac.uk/medical-education/other-studies/laterality/laterality-questionnaires

Morgan, M. J. & McManus, I. C. (1988). The relationship between brainedness and handedness. In F.C.Rose, R. Whurr, & M. Wyke (Eds.), *Aphasia* (pp. 85-130). London: Whurr Publishers.

Ranky, G. N., & Adamovich, S. (2010). Analysis of a commercial EEG device for the control of a robot arm. *Proceedings of the 2010 IEEE 36th Annual Northeast*. New York, NY.

Vandenberg, S. & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, *47*, 599-604.