2013

Predicting Cognitive Workload with Measures from Functional Near-Infrared Spectroscopy (fNIRS) and Heart Rate

John Duany
University of Central Florida

PREDICTING COGNITIVE WORKLOAD WITH MEASURES FROM FUNCTIONAL NEAR-INFRARED SPECTROSCOPY (fNIRS) AND HEART RATE

by

JOHN M. DUANY

A thesis submitted in partial fulfillment of the requirements
for the Honors in the Major Program in Psychology
in the College of Sciences
and in the Burnett Honors College
at the University of Central Florida
Orlando, Florida

Summer Term, 2013

Thesis Chair: Corey Bohil, Ph.D.

Abstract

The objective of this study was to assess low to high levels of Cognitive Workload by measuring heart rate and cortical blood flow in real-time. Four conditions were implemented into a within-subjects experimental design. Two conditions of difficulty and two conditions of trial order were used to illicit different levels of workload which will be analyzed with psychophysiological equipment. Functional Near-Infrared Spectroscopy (fNIRS) has become more prominent for measuring the blood oxygenation levels in the prefrontal cortex of individuals operating in hazardous work environments, students with learning disabilities, and in research for military training. This is due to the fNIR device being highly mobile, inexpensive, and able to produce a high-spatial resolution of the dorsolateral prefrontal cortex during executive functioning. Heart Rate will be measured by an Electrocardiogram, which will be used in concordance with fNIR oxygenation levels to predict if an individual is in a condition that produces low or high mental workload. Successfully utilizing heart rate and blood oxygenation data as predictors of cognitive workload may validate implementing multiple physiological devices together in real-time and may be a more accurate solution for preventing excessive workload.

Dedication

To Dad, Mom, David, Pop-Pop, and Grandma
For your unending love and encouragement

Acknowledgements

I would like to extend my gratitude to my advisor, Dr. Corey Bohil, for his guidance and dedication, and for allowing me the opportunity to perform my own research with the resources available in his laboratory. To Dr. Daniel McConnell, thank you for the last five years of your teaching and assistance, and for helping me advance my knowledge and skills with statistical analysis.   I would like to thank Dr. Enrique Ortiz for allowing me the opportunity to collaborate my ideas and research with him, along with his advice using fNIR Spectroscopy

My father, mother, brother, and grandparents deserve a special thank you for their tremendous support, encouragement, and prayers; all to help me learn life’s lessons, grow as a man, and succeed as a professional. Most importantly, I give thanks to Jesus Christ for providing me the willpower, courage, and hope to rise above life’s challenges

Table of Contents

List of Figures

Introduction

Analyzing cognitive performance and mental workload in real-time could be beneficial for a wide variety of critical operations. Mental workload (MWL) can be assessed using physiological measures as predictors, in addition to neuroimaging technology, for enhancing the safety and performance of individuals experiencing excessive workload in potentially hazardous working environments and time-critical military tasks.

The importance of assessing MWL in real-time is due to the difficulty of obtaining a fine-grained assessment of stress level if evaluated after a task has been completed. A rating scale can be used to assess workload after task completion, such as the NASA-Task Load Index (NASA-TLX), which consists of six scales (Hart and Lowell, 1988). The averages of these scales were weighted to illustrate how each factor may have an effect on the workload of a specific activity (Hart and Lowell, 1988). Subjects were instructed to specify which factors contributed least and most to the workload experienced during an experimental task. Requiring a person to remember their level of experienced workload is problematic, especially when trying to rate on scales from low to high. Memories can be easily altered or manipulated after a task of high MWL has been performed (Hart and Lowell, 1988). The possibility that stress and workload may be assessed during real-time rather than recall, could lead to a more reliable approach for observing cognitive workload without interfering with task performance. Allowing neurophysiological measurements to be collected in real-time will provide objective results and may improve insight into how real-time measures of heart rate and cortical blood flow could be used as predictors of stress and mental workload.

Functional Near-Infrared Imaging

Physiological recording, such as Electroencephalography (EEG) and Functional Near-Infrared Spectroscopy (fNIR), offers potential to read cortical activity in real-time as opposed to survey measures which rely on verbal reporting and prolonged recall. fNIR technology has assisted neuroscience researchers observe neural and cognitive features of executive functioning in addition to psychopathological and neurological deficits, in controlled and field experiment conditions. New techniques of assessing cortical activity have been developed using fNIR, allowing more diverse environments and scenarios to be implemented into studies to further understand executive functioning and brain activity in realistic environments (Bunce et al., 2007). fNIR has proven to possess more mobile capabilities than other neuroimaging equipment, including being safe to use portably in a moving military vehicle or during ambulatory care. It is inexpensive compared to fMRI and noninvasive to participants. It also has greater ecological validity when applied in laboratory or field conditions, especially when movement is required (Bunce et al., 2007).

According to modern neuroscience research, executive functioning is thought to occur in the prefrontal cortex, particularly Brodmann's Areas 9, 10, 45, and 46; and is involved with decision-making, sustained attention, problem solving, and working memory. During executive functioning, the prefrontal cortex requires glucose to be metabolized for energy in order to focus, sustain attention, or drive decision-making and problem solving. Oxygen is carried through the capillaries to the active areas of the brain and is released into the neural tissue. When oxygen enters the neural tissue, glial cells absorb the oxygen and metabolize glucose into energy, and then the blood in the activated areas becomes deoxygenated (Bunce et al., 2006).

Oxygenated and deoxygenated hemoglobin have elements observable in the near-infrared range of light, which allows researchers to view the change in hemoglobin concentration during neurovascular coupling. Neurovascular coupling is a process that stimulates the brain to increase arteriolar vasodilation. This increases blood flow and blood volume to the active areas of the brain, leading to the reduction of oxygen and the metabolization of glucose in the neural tissue (Bunce et al., 2006).

The increased blood flow to the brain, known as the hemodynamic response, will change in concentration from oxygenated hemoglobin to deoxygenated hemoglobin during brain activity. This process is observed in the dorsolateral prefrontal cortex by the NIR spectroscopy system. The fNIR sends photons at the near-infrared wavelength through the scalp then a small portion of those particles bounce back toward the light detectors of the spectroscopy system. Biological tissues (head, skin, skull, cerebral spinal fluid, brain) are transparent to light in the near-infrared range, between 700-900nm (optical window), due to water which absorbs a nominal amount of energy at *nir* wavelengths (Bunce et al., 2006). The sensors on the fNIR consist of four light sources which contain three built-in LEDs with set wavelengths at 730 nm, 805 nm, and 850 nm, along with ten detectors. Sixteen source-detector pairs (channels/optodes) were designed to provide imaging data from specific cortical areas involved with cognitive and executive functioning, the dorsolateral and inferior prefrontal frontal cortices (Bunce et al., 2006).

Combining fNIR with Heart Rate Data

Analyzing fNIR data, in conjunction with Electrocardiogram (ECG) data, may provide additional insight into MWL in an augmented cognition system or real-life operation during military training and time-critical working environments (Ayaz et al., 2011).

An augmented cognition program was established by the Navy and Defense Advanced Research Projects Agency (DARPA), concerning performance in education and training. The Naval Education and Training Command (NETC), the Human Performance Center (HPC), and Navy Personnel Development Command (NPDC) were developed to focus on individual training with the Integrated Learning Environment (ILE). Fleet Forces Command completed the rest of the training, team training with modeling and simulation based exercises (Nicholson et al., 2005). The objective was to optimize all levels of training and feedback from beginner to expert. fNIR and Electroencephalogram (EEG) were combined, among other physiological gauges, for regulation in a feedback loop. Increased heart rate was used as an indicator for arousal (Nicholson, et al., 2005). Results showed improvement in team performance by monitoring data and providing continuous feedback. In another example, an augmented cognition system assessed the user's cognitive state, read by physiological recording, which caused the system to adapt and change the task presented to the user in order to enhance the capabilities of the human in control (Stanney et al., 2009).

The Technical Integration Experiment (TIE) is an augmented cognition program developed by DARPA, used with naval training and education. DARPA used multiple physiological gauges for assessing cognitive state with level of workload. The TIE is a modified version of the Warship Commander Task, a simulator for naval airbase training developed by

DARPA that can activate numerous facets of cognition and emulate environments that require operational warfighters in tactical command centers to perform complicated decision-making (St. John et al., 2004). A cognitive performance measurement study was developed by DARPA, using fNIR to assess a participant's cognitive workload by observing changes in blood oxygenation levels. The rate of change in blood oxygenation was significantly responsive to workload variations across the left and right hemispheres (Bunce et al., 2007).

DARPA adapted the TIE to make the gauges more sensitive to participants, measuring six dimensions of workload, which included using fNIR and EEG gauges. EEG, fNIR, and event related potentials (ERP's) were used concurrently to produce high spatial resolution in time intervals for detecting changes in workload which allowed the task to adapt for the participants cognitive state to reach an acceptable level of performance (St. John et al., 2004).

In a system with integration of fNIR, EEG, and an ECG, a controller analysis was needed to make adjustments during tasks which were causing mental overload. The system provided feedback based on the user's condition, which can be read by an EEG. Beta, alpha, and theta bandwidths were used as indicators by the EEG which adapted the task allowing the participant to meet task demands; the controller analysis monitored the feedback loop and adjusted the task according to these indicators (Stanney et al., 2009). A Control Framework determined the best conditions for adapting fNIR sensor data and keeping cognitive workload at a desired range by employing a set of decision rules. Distractors and lack of attention was alleviated in a control framework with attenuators, time reactions, and detectors which measured whether the participant was meeting cognitive demands (Stanney et al., 2009).

Our study was designed to see how well we can measure small changes in mental workload in real-time. Each participant in the experimental group was engaged in a cognitive task, specifically counting backwards. The task was split into two parts, easy and hard-counting trials were performed alternately, and then easy and hard-counting trials were presented in random order. All participants wore an fNIR sensor and ECG electrodes attached to the right wrist and directly below both ankles to measure heart rate. The difference of mental workload elicited by the difficult tasks should be observed through changes in fNIR and heart rate data. If fNIR and ECG can be used as predictors of mental workload, the combination of these measures could provide a more accurate assessment of workload in real-time.

We sought to design an algorithm that could predict cognitive workload state using the multiple sensor technologies, fNIR and ECG, similarly to the gauges being used in DARPA's Technical Integration Experiment and Integrated Learning Environment. Analyzing cognitive performance and mental workload in real-time, could provide insight into how researchers can enhance safety and performance by avoiding excessive workload of individuals and teams working in high-risk environments and time-critical military operations.

Methods

Participants

Ten participants volunteered for the study via the University's SONA research participant pool. The participants were university students with no restrictions placed on age, gender, or ethnicity. All participants must have been enrolled in an undergraduate psychology course for access to sign-up for the study. IRB approval was required for the use of the University students as participants.

Design

All subjects completed four conditions in a within-subject experimental design. Four conditions were developed with a total of 40 trials; two difficulty conditions, easy and hard, and two trial order conditions, alternating order and random order. All participants completed the easy-condition (counting backward by 7s) and the hard-condition (counting backward by 7s). An alternating counting task (task 1) and a random order task (task 2) was given to nine participants. The data of one participant (subject 1) was omitted for having outlier data.

Materials

The fNIR System 100 was developed by Drexel University and distributed by BIOPAC Systems, INC. The electrocardiogram was also distributed by BIOPAC Systems, INC. The fNIR System 100 was connected to a Hewitt-Packard laptop via USB, the ECG was connected to a Dell desktop computer via USB. A second laptop (TOSHIBA) was used to track all data from the counting tasks.

Procedure

The ECG was connected to a desktop computer, while the fNIR was located on a cart and connected to a laptop. Once both programs were up and running on each computer, the electrodes for the ECG were attached to the participant starting with the positive and neutral (red/black) charged electrodes on the feet, followed by a negative electrode on the right wrist. The participant was asked to sit completely still while baseline data for the ECG was recorded and visually verified as being clean. The fNIR was secured onto the forehead and strapped to the back of the head, then fastened with an elastic non-stick tape which also helped block out ambient light. A ten second baseline was recorded for the fNIR before the experiment began.

The participant was instructed to perform a cognitive task which involved counting down by sevens out loud from a two or three-digit number. A marker on the fNIR record and a marker on the ECG record were placed simultaneously to signify the beginning of the study. Forty trials were performed divided into an easy-condition (counting down by sevens from a two-digit number) and hard-condition (counting down by sevens from a three-digit number), separated into two sets of trials of twenty each. The participant was given a two or three-digit number and began counting after the researcher said “begin”. The participant counted for ten seconds, and then the researcher said “stop”, and waited ten seconds before beginning the next trial.

Results

Counting Task Performance

The average quantity of numbers counted for the easy-counting condition were higher than the numbers counted in the hard-counting condition in the alternating trial order, as depicted in Figure 1, (easy: M = 7.95, hard: M = 6.3). The average numbers counted for the easy-condition was higher than the numbers counted in the hard-condition, in the random trial order, see Figure 2, (easy : M = 8.125, hard: M = 6.82). The median and mode were calculated for easy and hard-counting conditions in the alternating and random order trials; alternating easy-condition (*MDN* = 8, *MD* = 8), alternating hard-condition (*MDN* = 5, *MD* = 4), and random order easy-condition (*MDN* = 7, *MD* = 7), random order hard-condition (*MDN* = *6*, *MD* = 5).

A two-sample t-test was conducted to compare performance within two levels of the alternating counting trials and random order counting trials. There was a statistically significant difference in the amount of numbers counted between easy and hard-counting conditions in the alternating order trials**,** $t(8)=2.58$, $p=.016$. These results suggested the hard trials required greater cognitive workload than the easy trials for the alternating counting order. The average numbers counted by trial, in random order, also showed a significant difference between the easy and hard-counting conditions, $t(8)=4.35$, $p=.001$.

The number of errors made during the easy and hard-counting conditions in the alternating order trials were summed across subjects 2 through 8. The predicted result of having more errors in the hard-counting condition than the easy-counting condition is shown in Figure 1 (top; *total errors for alternating counting made (easy)* = 15, *total for errors random order (hard) made* = 20). The number of errors made during the easy and hard-conditions from the random order trials were also summed across subjects 2 through 8, in contrast to the alternating order data, the easy-condition produced more errors (*total errors made* = 18) than the hard-condition (*total errors made* = 17), see Figure 1 (bottom). A t-test was performed to compare number of errors within each subject for the easy and hard-conditions in the alternating counting trials and then calculated separately for the random order trials. Both appeared not to be statistically significant, alternating order errors, $t(8)$ =-.887, p=.200, and random order errors, $t(8)$ =.186, p=.436. An additional t-test was used to calculate the errors as a percent of numbers counted, which turned out not to have a significant difference in error rates by percentage or by overall error tallies, $t(8)$=.186, p=.429.

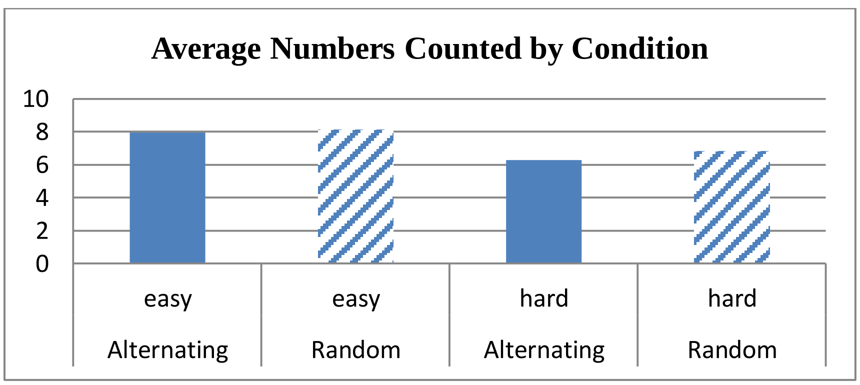
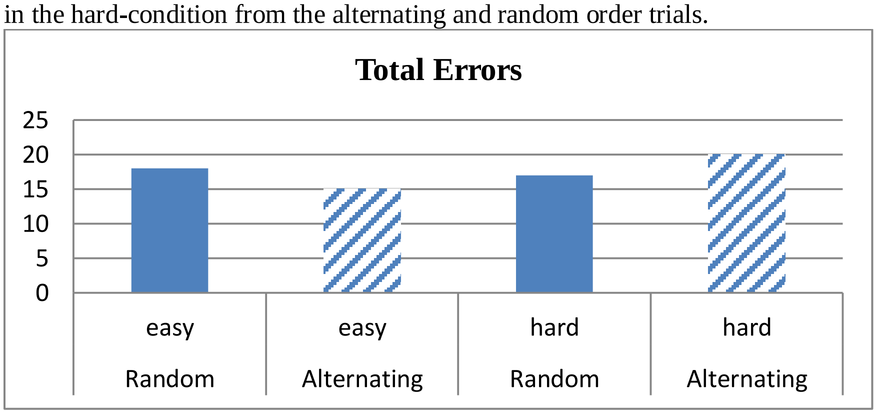


Figure 1 Top: Average numbers counted in the easy-condition from the alternating and random order trials were compared to the average numbers counted in the hard-condition from the alternating and random order trials. Bottom: Total errors made in the easy-condition from the alternating and random order trials were compared to the total errors made in the hard-condition from the alternating and random order trials.



Counting Task Performance ANOVA

A 2x2 within-subjects ANOVA was conducted to analyze counting performance differences between difficulty and trial order conditions. Collapsed over order, the difference between the easy and hard-conditions was approaching significance, $F(1, 8)= 4.111$, $p=.077$, *partial* $\eta^2=.339$. Collapsed over difficulty, a significant difference was observed between alternating and random trial order, $F(1, 8)=18.118$, $p=.003$, *partial* $\eta^2=.694$. The interaction of difficulty and trial order is not significant, $F(1, 8)=.114$, $p=.744$, *partial* $\eta^2=.014$.

fNIR Results

The change from baseline oxygenation level was higher for the easy-condition, (M = .516, SD = 0.253), than the hard-condition, (M = 0.386, SD = .275). A t-test compared the easy and hard-condition means, $t(5)$ =.726, p=.25, and did not find a difference large enough to be statistically significant. None of the differences were significant except between the easy and hard-conditions in channels 1 and 16. Oxygenation differences between easy and hard-conditions in Channel 1 were almost significant, $t(5)$=.077, p=.076, and Channel 16 did show a significant difference, $t(5)$=.03, p=.029. A t-test was performed with the easy-condition for the alternating counting trials comparing the average oxygenation levels of the left optodes (1-8) to the optodes on the right (9-16), for six participants. The difference seen between optodes 1-8 and 9-16 was in the direction of being significant, $t(5)$=-1.56, p=.089, and no significant difference was found in the hard-condition, $t(5)$=-1.182, p=.145. Eleven of the 16 Channels depicted higher levels of oxygenation from the hard trials over the easy trials (Figure 2). Although the effect was not enough to be statistically reliable, it did appear to be fairly consistent across measurement locations.

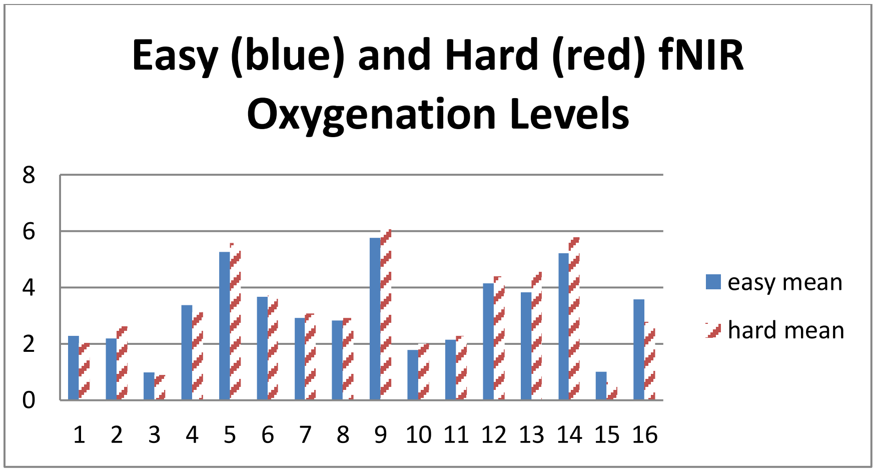
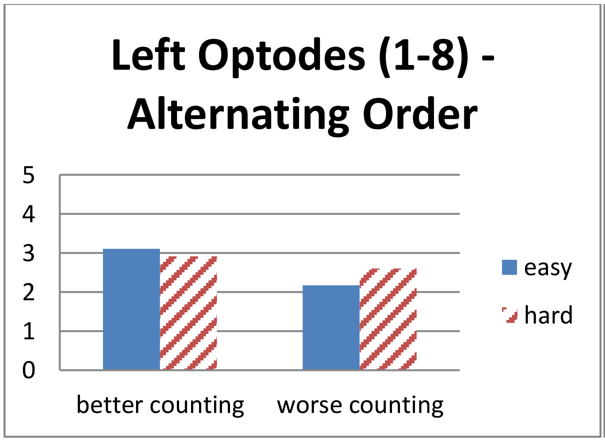


Figure 2: Increase over baseline fNIR values for each channel (optode) location.

Left optodes (1-8) and right optodes (9-16) were compared by analyzing the averages of the oxygenation levels for the participants, grouped together as better counters, against the participants who were grouped together as worse counters. For the alternating and random order counting tasks, the better counters (subject 3, subject 7, subject 9 for alternating; subject 3, subject 6, subject 7 for random) showed higher oxygenation levels than the poorer counters in both the easy and hard trials on the left optodes. (Figure 3)



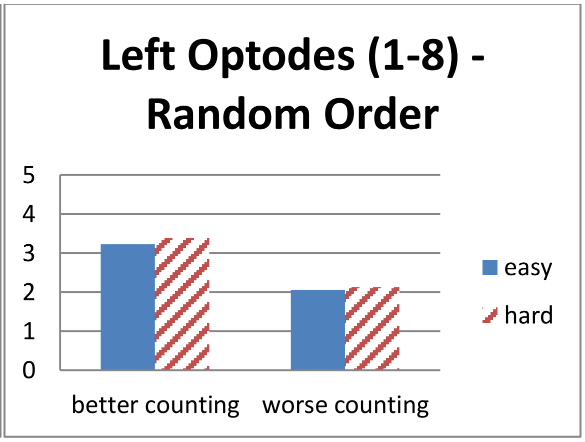


Figure 3: Left: Left fNIR Optodes depicting oxygenation levels of the best and worst counters from alternating order. Right: Left fNIR Optodes depicting oxygenation levels of the best and worst counters from random order.

The fNIR indicated lower oxygenation levels on the right optodes (figure 4 below: left and right)

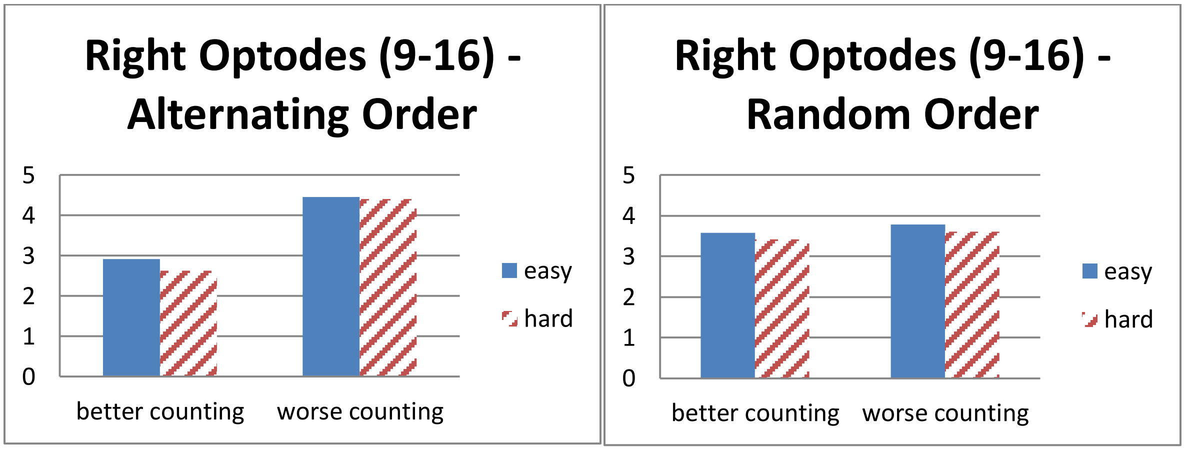


Figure 4 Left: Right fNIR Optodes depicting oxygenation levels of the best and worst counters from alternating order. Right: Right fNIR Optodes depicting oxygenation levels of the best and worst counters from random order.

With all the channels collectively measured from the alternating counting trials, the worse counters appeared to require higher levels of oxygenation than the better counters (Figure 5: left), though the better counters appeared to require more oxygenation from the random order trials (Figure 5: right).

Figure 5: Average fNIR change from baseline by counting performance. Left: alternating order trials, right: random order trials.

fNIR ANOVA

A 2x2 within-subjects ANOVA was conducted to analyze the difference in oxygenation levels with statistics from two measures of difficulty (easy and hard), trial order (alternating and random), and the interaction of difficulty and trial order. The easy alternating trials and the easy random order trials were paired against the hard alternating trials and the hard random order trials. Difference of difficulty did not have an effect on the oxygenation levels, $F(1,5)=.000$, $p=.984$, *partial* $\eta^2=.000$. No significant difference was noted for trial order (alternating against random), $F(1,5)=.126$, $p=.737$, *partial* $\eta^2=.025$, and no interaction was found between difficulty and task order, $F(1,5)=.029$, $p=.871$, *partial* $\eta^2=.006$.

Electrocardiogram Results

Figure 6 displays average heart rate (task BPM – baseline BPM) for the four experimental conditions. A 2x2 within-subjects ANOVA was performed to measure the difference between the easy and hard-conditions and across alternating and random order trials. A small main effect was found from difficulty, $F(1, 5)=.216$, $p=.661$, *partial* $\eta^2=.041$. There was no significant difference between alternating and random order trials, $F(1,5)=.1.235$, $p=.317$, *partial* $\eta^2=.198$. The interaction between difficulty (easy versus hard) and trial order (random order versus alternating) was not significant, $F(1,5)=3.037$, $p=.142$, *partial* $\eta^2=.378$.

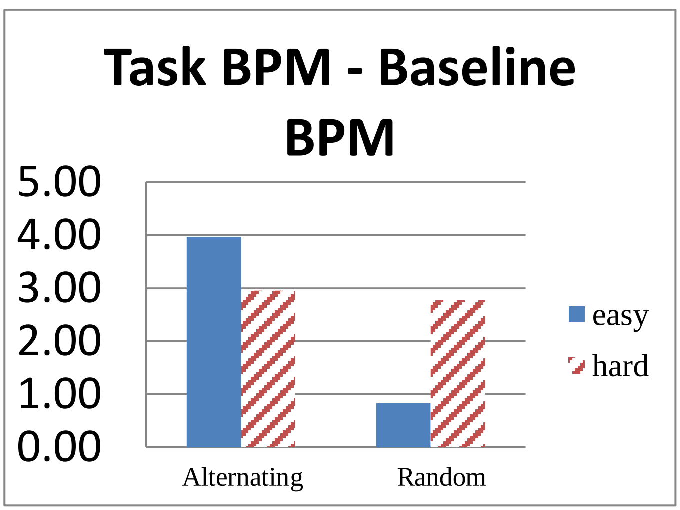


Figure 6: Baseline BPMs deduced from the task BPMs to analyze the difference between easy and hard-conditions within alternating and random order trials.

Focusing on the "random order" trials only, there was a larger increase in BPM in the hard-condition (*M* = 2.771, *SD* = 6.503) over the easy-condition (*M* = .0826, SD = 4.256). The trend was in the predicted direction, but was not significant, *t(5*)=,-1.146, *p* =.15.

For the alternating trial conditions, there was an increase in BPM in the easy-condition (*M* = 3.972, *SD* = 3.613) over the hard-condition (*M* = 2.960, *SD* = 3.743). The alternating task t-test indicated that the easy and hard difficulties were approaching a significant difference in heart rate, *t(5*)=1.35 *p*=.117. Both easy-conditions were paired from the alternating and random order trials. A t-test was conducted and found a significant difference between both easy-conditions in the alternating and random order trials, *t(5*)=3.07, *p*=.014. No differences were reported for the hard-conditions from the alternating and random order trials, *t(5*)=-.085, *p*=.468.

Heart Rate Variability

Variability based on BPM – Baseline was highest for the hard-condition in the random order trials (*SD* = 9.791), while the easy-condition in the random order trials had the lowest variability (*SD* = 5.341). Average alternating variability was slightly greater in the hard-condition (*SD* = 6.941), than the easy-condition (*SD* = 6.089). The difference in variability between difficulty level in the alternating task approached significance, *t*(5)=-.833, *p*=.221. The variability between easy and hard in the random order trials was closer to being significant, *t*(5)=-1.39, *p*=.125. There was not a significant difference of variability in relation to the interaction between difficulty and trial order. *t(5*) =-.584, *p*=.332.

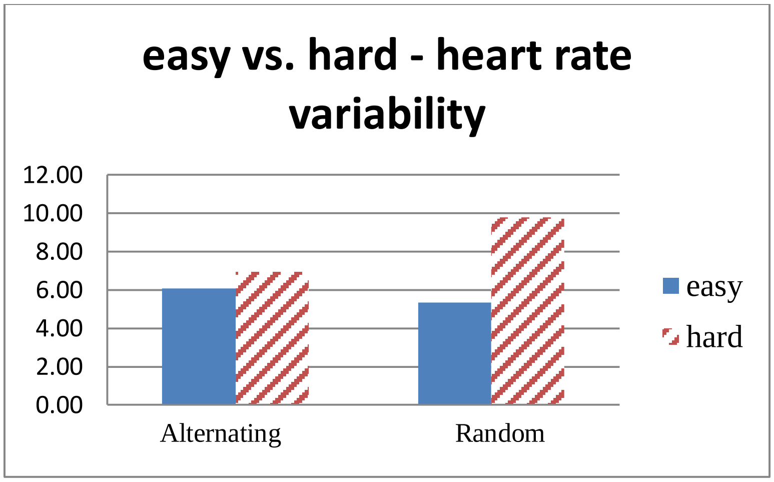


Figure 7: Variability between the easy and hard-conditions with the alternating and random conditions.

A 2x2 within-subjects ANOVA was conducted to analyze the difference in variability from BPM – Baseline statistics with two measures of difficulty (easy and hard), task order (alternating and random order), and the interaction of difficulty and task order. The difference of difficulty was not significant, $F(1,5)=1.755$, $p=.243$, *partial* $\eta^2=.260$. Task order effect size was progressing toward the direction of significance, $F(1,5)=.332$, $p=.589$, *partial* $\eta^2 = .062$; and a weak interaction was found between difficulty and task order effects, $F(1,5)=1.358$, $p=.296$, *partial* $\eta^2=.214$.

MANOVA

A within-subjects repeated measures Multivariate Analysis of Variance (MANOVA) was performed, analyzing the interaction and differences of difficulty and trial order between the ECG heart rates and the fNIR oxygenation levels. Difficulty - $F(1,5)=1.65$, $p=.3$, *partial* $\eta^2=.452$ Trial Order - $F(1,5)=.45$, $p=.67$, *partial* $\eta^2=.184$, Interaction of difficulty and trial order - $F(1,5)=1.378$, $p=.351$, *partial* $\eta^2=.408$.

Correlated Data

fNIR oxygenation levels and ECG heart rate were evaluated for a correlation. A correlation was not found between the fNIR oxygenation levels and heart rate, as depicted in figure 8, $r = -0.182$. Results from the correlation were contrary to our prediction; high heart rate could not be interrelated with high oxygenation levels of the same task for each participant.

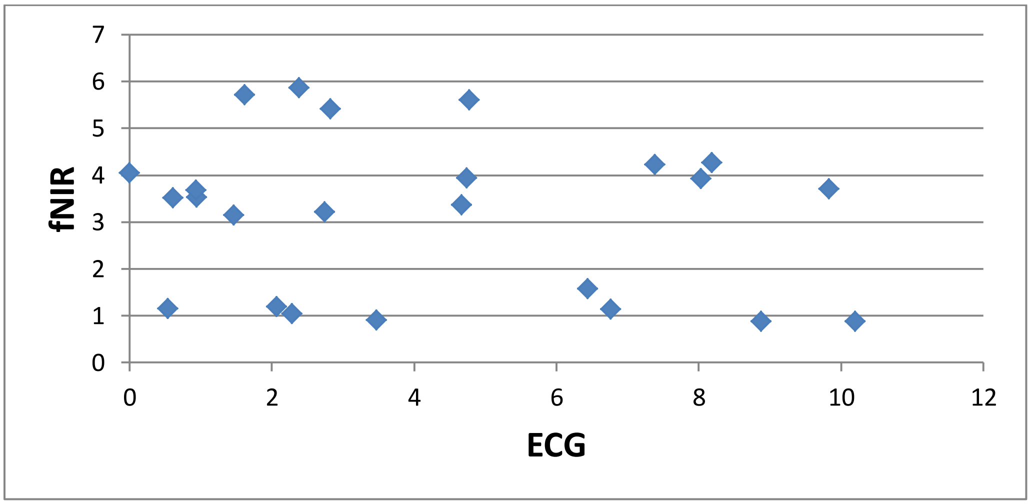


Figure 8: fNIR oxygenation levels and ECG heart rate correlated on a scatterplot.

Predicting Workload from Heart Rate and Brain Activity

Based on fNIR data, the average oxygenation levels were observed from twenty alternating trials and twenty random order trials. The oxygenation means of the alternating trials were used to guess whether the random ordered trials were in a high or low workload condition, based on deviations from the easy or hard means. For example, if the deviation of a hard random trial mean from an easy or hard-alternating trial mean is greater than the deviation of an easy random trial mean from an easy or hard alternating trial mean, then we can guess that a trial was in the easy-condition. Only one instance was not predicted correctly, the rest were correctly predicted with a 92% average.

Discussion

Counting Task

The participants completed both easy and hard mental arithmetic tasks while their heart rate and Blood Oxygenation Level-Dependent (BOLD) signal in the prefrontal cortex were monitored. In the easy-condition, participants counted down by sevens (aloud) from two-digit numbers. In the hard-condition, participants counted down by sevens (aloud) from three-digit numbers. Forty trials were separated into two types of trial sequence. The first twenty trials were counted in alternating order, i.e. easy/hard/easy/hard; the last twenty trials were counted in random order. These two conditions were intended to produce different levels of mental workload. Our goal was to observe measurable changes in BPMs and blood oxygenation levels and use those changes to predict whether a participant is in a high or low workload condition. The counting tasks provided the expected results: more numbers counted in the easy-condition than the hard-condition for both alternate and random trial order.

Electrocardiogram

Higher heart rate was found for the easy-counting condition compared to the hard-counting condition within the alternating trials. The participants may have had to walk across campus to get to the study, which could explain the overall higher heart rate averages from the alternating trials. Lower overall BPMs were found in the random order task though the hard trials had higher heart rate averages than the easy trials. The Electrocardiogram data was limited to six participants, which may have led to unexpected results with the easy and hard-conditions.

The anticipation to count may have also led to overall higher heart rate in the alternating counting task. The participant may have had anxiety from engaging in counting under a time limit, or felt ill-prepared to engage in a cognitive activity with a duration lasting about thirty minutes. A questionnaire was not handed out after the study was performed but would have been helpful in assessing some of the difficulties the participants might have had during each task and how any previous activities could have affected their heart rate before beginning the study.

fNIR Results

The blood oxygenation levels were higher on the left side of the prefrontal cortex, channels 1-8, in the alternating and random ordered trials, when compared to the right side of the prefrontal cortex, perhaps due to physiological differences. To speculate, the lateralization of brain function to left and right-handedness may have played a role when observing the difference of oxygenation levels from best and worst counters. Specific areas on the left side of the dorsolateral prefrontal cortex may be more active than the right side for more mathematically acclimated minds. The right channels (9-16) showed higher oxygenation levels for the worse counters. This could denote that more mental energy was being manifested in structural areas of the right prefrontal cortex. The low level of significance between the easy and hard-condition is most likely due to the hard-condition not being difficult enough. Instead of doing a counting task, the hard-condition could have required calculations such as multiplication and division or a Fibonacci sequence. The hard-condition was too similar to the easy-condition and needed to require higher levels of workload to produce notable differences.

Predicting Workload

We were successfully able to predict whether than random trials were in a high or low workload condition, based on the deviation of the alternating trials from the means of the random order trials, with 92% accuracy. A significant difference may exist between random and alternating trials by looking at results from an ANOVA, though the 92% average is likely inflated due to that difference.

A correlation test indicated that the heart rate data did not correlate with the oxygenation levels. We were expecting to see elevated heart rate paired with high oxygenation levels for the same participant in a task of high workload. A repeated measures within-subjects MANOVA was performed with heart rate levels and fNIR oxygenation levels and found no interaction between the two physiological measures. No significance was found for difficulty, trial order, or the interaction of difficultly and trial order between ECG and fNIR data.

The results found by DARPA suggest that the Technical Integration Experiment could employ multiple psychophysiological gauges to accurately assess changes in cognitive workload during specific command and control tasks. The objective for part of the experiment was to integrate the gauges into a system that demonstrates real-time manipulation of cognitive states as the structure of the augmented cognition system (St. John et al., 2004). We were able to employ physiological gauges to assess cognitive workload, similar to the TIE, although an augmented cognition was not developed to regulate feedback to the user.

Several changes could have been implemented to the study to make a complex, realistic scenario and environment, along with modifying the counting task to a task which more suitably produces high workload. These changes would have provided stronger differences between heart rate and oxygenation levels in each condition of difficulty and trial order. Diverging from counting and instead using simulation training, such as airspace monitoring from the Warship Commander task, a driving simulator, a non-violent Source engine game, or a dynamic environment that simulates operations being performed on the International Space Station; might have produced distinct differences between low to high levels of workload. The counting task did show promise for driving a considerable difference between workload levels. With an increase of ten or more participants, there may have been larger effect sizes with significant differences between the difficulty levels and trial order from fNIR and ECG data, in addition to a correlation of increased oxygenation levels with elevated heart rate in a task requiring high mental workload.

References

Ayaz, H., Shewokis, P.A., Bunce, S., Curtin, A., Izzetoglu, K., Onaral, B., Willems, B., in press. (2011). Optical brain monitoring for operator training and mental workload assessment. *NeuroImage* 59, 36-47.

Bunce, Scott C., Izetogu, K., Izzetogu, M., Onaral, B., Poirrezaeo, K. (2006). An Emerging Neuroimaging Modality. *IEEE Engineering In Medicine and Biology Magazine* July/August, 54-60.

Bunce, Scott C., Izetogu, K., Izzetogu, M., Onaral, B., Poirrezaeo, K. (2007). Assessing Cognitive Activity in Real-Life Situations. *IEEE Engineering In Medicine and Biology Magazine* July/August, 38-45.

Hart, Sandra G., and Lowell E. Staveland. (2008) "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." *Advances in Psychology* 52: 139-183

Henellius, A., Hirvonen, K., Holm, A., Korpela, J., Muller, K. (2009). Mental Workload Classification using Heart Rate Metrics. *31st Annual International Conference of IEEE EMBS.*

Huppert, T.J., Hoge, R.D., Diamond, S.G., Franceschini, M.A., Boas, D.A. (2006). A temporal comparison of BOLD, ASL, and NIRS hemodynamic responses to motor stimuli in adult humans. *NeuroImage* 29, 368–382.

Nicholson D., Lackey, S., LT Arnold, R., Scott, K. (2005). Augmented Cognition Technologies Applied to Training: A Roadmap for the Future.

Stanney, K.M., Schmorrow, D.D., Johnston, M., Fuchs, S., Jones, D., Hale, K.S., Ahmad, A., Young, P. *Reviews of Human Factors and Ergonomics, 5*: 195

St. John, M., Kobus, D. A.,. Morrison, J.G., Schmorrow, D.D. (2004). *Overview of the DARPA* Augmented Cognition Technical Integration Experiment. *International Journal of Human-Computer Interaction, 17(2)*:131-149.