



**SMALL-MUSCLE MOVEMENT AND ITS EFFECT ON COGNITIVE
PERFORMANCE
THESIS**

Timothy R. Meyer, Captain, USAF

AFIT-ENG-MS-17-M-053

**DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY**

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government. This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

AFIT-ENG-MS-17-M-053

SMALL-MUSCLE MOVEMENT AND ITS EFFECT ON COGNITIVE
PERFORMANCE

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Computer Science

Timothy R. Meyer, BS

Captain, USAF

March 2017

DISTRIBUTION STATEMENT A.
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AFIT-ENG-MS-17-M-053

SMALL-MUSCLE MOVEMENT AND ITS EFFECT ON COGNITIVE
PERFORMANCE

THESIS

Timothy R. Meyer, BS
Captain, USAF

Committee Membership:

Dr. Brett Borghetti
Chair

Maj. Jason Bindewald, PhD
Member

Mr. Justin Estepp
Member

Abstract

The research community has developed methods to be able to detect the cognitive state of operators within the military, and are continuing to improve on those methods. However, that detection is of no value without the ability to positively affect an operator's cognitive state. This thesis study provides the ground work for exploring ways to improve the cognitive state of an operator as needed. Small-muscle warm-up, such as hand and finger movement, is widely unexplored for its effects on cognitive performance. The experiment investigates the use of a hand and finger warm-up activity and its effects on reaction time, categorization, impulse control, episodic memory, and visual spatial (VS) working memory. Sixteen participants each completed five cognitive measurement activities under two different conditions: preceded by a hand and finger warm-up activity and also preceded by a relaxed state activity. Results showed that hand and finger activity significantly increased reaction time in three measures of cognitive performance, the largest increase being a reaction time of 8 milliseconds faster in the treatment condition than under the control condition. The analysis also suggested an increase in episodic memory. In addition to the effect of the warm-up, the pace with which participants conducted the warm-up has a significantly positive correlation coefficient of 0.54 with regards to the size of the cognitive effect for reaction time. The findings of this study show that small muscular movements do have a significant effect on cognitive functions.

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Dr. Brett Borghetti, for his guidance and support throughout the course of this thesis effort. He took great care and effort during each step in the process to ensure I had what I needed to succeed. Dr. Borghetti was always willing to provide his insight and experience to whatever challenges arose. I consider it very valuable the many skills, techniques, and the knowledge that I have had the privilege to glean from him.

Timothy R. Meyer

Table of Contents

	Page
Abstract.....	iv
Acknowledgments.....	v
Table of Contents.....	vi
List of Figures.....	ix
List of Tables.....	x
I. Introduction.....	1
General Issue.....	1
Research Questions and Hypotheses.....	5
Approach.....	6
Assumptions/Limitations.....	7
Key Findings.....	8
Thesis Structure.....	9
II. Literature Review.....	11
Chapter Overview.....	11
Research Gap.....	11
Cognitive Measurement Activities and Associated Cognitive Functions.....	13
Test Designs.....	18
Results.....	21
Areas of Uncertainty.....	23
Summary.....	24

III. Methodology	26
Chapter Overview.....	26
Research Questions	26
Hypotheses	28
Dependent Measures	29
Cognitive Measurement Activities	32
Treatment Conditions	39
Hidden Factors.....	43
Experiment Sequence	43
Assumptions	45
Data Collection.....	46
Treatment – Hand/Finger Movement T10 (Reaction time):.....	56
Summary.....	57
IV. Analysis and Results.....	58
Chapter Overview.....	58
Experiment Timing and Latency Considerations	59
Reaction time Results & Discussion – Reaction Time Test (T1).....	60
Categorization Results & Discussion – Oddball and Go/No Go (T2).....	73
Inhibitory/Impulse Control Results & Discussion - Go/NoGo (T3)	74
Episodic Memory Results & Discussion – Word Recall (T4)	75
Visual spatial working memory Results & Discussion – Dot Matrix (T5)	77
Summary.....	79
V. Conclusions and Recommendations	81

Chapter Overview.....	81
Conclusions of Research	81
Significance of Research	83
Recommendations for Action & Future Research.....	83
Summary.....	88
Appendix A: Reaction time Results.....	89
Appendix B: Learning Effect Individual Analysis	93
Appendix C: Word Recall List	95
Appendix D: Pre and Post-Experiment Questionnaires.....	97
Appendix E: Word Recall Time Score Pseudo Code	100
Appendix F: Approval Letter.....	101
Bibliography	103

List of Figures

	Page
Figure 1: Dot-Matrix with grid	15
Figure 2: Reaction Time Test Example	33
Figure 3: OddBall and Go/NoGo Display	34
Figure 4: Six Gabor Patch Variations	35
Figure 5: Word Recall Display	36
Figure 6: Dot Matrix Equation.....	37
Figure 7: Dot Grid Display	38
Figure 8: Treatment Activity	40
Figure 9: Time Series of Correct vs Incorrect Oddball Responses.....	50
Figure 10: Time of Entry Score Function.....	55
Figure 11: Effect Size of Treatment.....	64
Figure 12: Reaction Time Averages over Time.....	65
Figure 13: Effect Size of Treatment for Age When Started Video Games	67
Figure 14: Effect Size of Treatment for Years Video Games Played.....	68
Figure 15: Effect Size of Treatment for Age	69
Figure 16: Effect Size of Treatment for Exercise Frequency	70

List of Tables

	Page
Table 1: Cognitive Effects and Treatments Used	21
Table 2: Research Questions.....	27
Table 3: Research Hypotheses	29
Table 4: Dot Matrix Series of Sets.....	39
Table 5: Experiment Sequence	44
Table 6: Example Go/NoGo Confusion matrix	51
Table 7: Example Confusion Matrix for Dot Matrix Equation.....	56
Table 8: Result of Means	59
Table 9: Scoring Approaches for Word Recall.....	76
Table 10: Reaction time Results by Participant.....	90
Table 11: Hypothesis's for Two-Sided t -tests	91
Table 12: Learning Effect Results by Participant.....	94
Table 13: Word Recall Lists	95

SMALL-MUSCLE MOVEMENT AND ITS EFFECT ON COGNITIVE PERFORMANCE

I. Introduction

General Issue

Cognitive performance an important factor in the workforce and military. Murnane et al. studied how cognitive ability can be used to estimate an individual's wages, stating this is possible to do because of the increased occupational demand for cognitive skills [1]. However, even if workers or military personnel meet the required cognitive skills for a particular position, the conditions of the job may induce fatigue or other distractors that can reduce cognitive performance. Thus researchers are looking for methods for improving cognitive performance. Exercise has long been studied for its positive effects on cognitive performance [2]. More recently and specific to one occupation, the surgical field has been conducting several studies based around small-muscle warm-up before an operation and the associated cognitive effects [3]–[7]. Small-muscle warm-up has been occurring for years, if not decades, among the video gaming community, though it remains largely unstudied. In a community board on IGN.com, players were asked how they warm-up their fingers before a game. Responses ranged from playing the piano to doing a practice round where the score doesn't matter to moving fingers in the air [8]. The one thing in common between the surgical studies and the video gamers is that the warm-ups include small-motor movement prior the task. The purpose of this thesis is to explore whether this small-muscle movement prior to a task

does improve cognitive performance.

The US Military employs many different types of operators including office workers, surgeons, pilots, and air battle managers. These operators share one thing in common; they use small-muscle movement to complete their tasks. However, to our knowledge the surgical field has published research on the benefits of small-muscle warm-up. Many small-muscle performance based studies have been and are being conducted within the surgical domain. Several of these completed studies have shown significant impact on the performance of surgeons in the Operating Room which includes the reduction of errors made [6]. But what other operator fields could benefit from warm-up? In the realm of Remotely Piloted Aircraft (RPA), a study done by Tvaryanas et al. revealed that on average, for every 100,000 flight hours by the Navy/Marine's RQ-2 Pioneer, 334 mishaps occurred [9]. This can be compared to a rate of 1 mishap per 100,000 flight hours for general aviation. Of the 334 mishaps, 32.9% were due to cognitive and psycho-behavioral factors, which include an individual's attention and complacency respectively. The other 67.9% is made up of factors outside the individual's control such as organizational processes or physical environment [9]. Clearly, RPA operations could be one of many areas that could benefit from methods of improving cognitive performance.

The current methods for improving an operator's cognitive performance are limited by our understanding of exactly what is causing a cognitive effect and the use of performance metrics which are too field specific. Research that demonstrates this

confusion as to the root cause of cognitive performance changes is gum chewing. This heavily studied area is plagued by many complexities, such as when to chew, that limit our understanding and have led to studies that show seemingly opposite results; some see a cognitive improvement after gum chewing, while others do not [10]. However, gum chewing studies do tend to focus on clearly defined cognitive functions, which is something not heavily found in surgical studies. Surgical studies have used warm-up methods from cell phone gaming to tasks like virtual reality scenarios, however, most methods for gauging performance after using said warm-up methods are specific to the surgical field and are therefore too narrow for application in other fields [3], [6]. We are left with this idea that warm-up is beneficial to improving cognitive functions, but it is hard to say with any definitiveness what the exact triggers are and which cognitive functions are affected. The area of small-muscle movement and the associated effects on cognitive performance is ill-defined. However, we do know that cognitive improvements can help reduce operator errors. Lendvay et al. claim a virtual reality robotic surgery warm-up, which is operated via hands and fingers, reduced errors by 33% [6]. Though ill-defined, there is value in better understanding the relationship between small-muscle movement and cognitive performance.

To better define the triggers and the affected cognitive functions, this thesis removes any similarities between warm-up and the task to be performed. Using an unrelated warm-up, we are able to look at the effects on cognitive performance with greater precision. Cognitive performance itself is usually a combination of several cognitive functions being measured. These cognitive functions may include but are not

limited to: attention, perception, memory, automaticity, language [11]. Within the Air Force, RPA pilots face many different mental strains that affect their cognitive functions including lengthy periods of time at a workstation environment producing sedentary behavior [12]. It is well known from large muscle exercise studies that activity can improve cognitive functions such as reaction time, visual discrimination, or short-term memory [13]. Such cognitive functions are pivotal within a RPA pilot's job, and while a pilot may not be able to stop and exercise during these prolonged periods of screen time, they may be able to do a quick small-muscle activity that would help improve their cognitive functions.

This research investigates the effect of small-muscle activity on various cognitive functions: reaction time, categorization, impulse control, episodic memory, and visual spatial (VS) working memory. Reaction time improves if an individual is able to respond to a stimulus faster. Categorization improvement is the individual's ability to better identify the correct stimuli. Impulse control improvement is an individual's ability to more often prevent a habitual response. Episodic memory improvement is an individual's ability to increase the number of event details experienced after a significant break in time. VS working memory improvement is an individual's ability to better process spatial or positional data while performing multiple interwoven tasks requiring the temporary storage and manipulation of information. Surgical studies have used various activities to improve such cognitive functions, but these warm-up activities are often related to the surgical task such as a custom Nintendo Wii game [14], or a Virtual Reality Robot [6], both specifically designed to simulate surgical motions. This relationship to the task

prevents any of these surgical studies from concluding that the small-muscle movement during the warm-up activity was solely responsible for the improvement in cognitive functions. Because the small-muscle activity used in this thesis study has no relation to the tasks used to measure each cognitive function, the findings are more broadly applicable and the cognitive improvement activities are not limited to RPA pilots, but can be adapted for the benefit of any military operator [15].

In the following section we describe a set of research questions, which explore whether a non-task specific small-muscle warm-up exercise can improve general cognitive performance.

Research Questions and Hypotheses

- Main question: Does small-muscle movement before a task have an effect on the cognitive performance of an individual during the task? Hypothesis: Cognitive performance is increased after hand/finger movement.
- Does reaction time improve with prior hand/finger movement? Hypothesis: Reaction time is faster after hand/finger movement.
- Do reaction time improvement gains due to hand/finger movement diminish over time? Hypothesis: Reaction time is faster after hand/finger movement, but the improvement diminishes over time.
- Does categorization improve with prior hand/finger movement? Hypothesis:

Categorization is more accurate after hand/finger movement.

- Does inhibitory/impulse control (executive control) improve with prior hand/finger movement? Hypothesis: Inhibitory/impulse control is more accurate after hand/finger movement.
- Does episodic memory accuracy improve with prior hand/finger movement? Hypothesis: Episodic Memory is more accurate after hand/finger movement.
- Does episodic memory recall Speed improve with prior hand/finger movement? Hypothesis: Episodic Memory Recall Speed is faster after hand/finger movement.
- Does visual spatial working memory improve with prior hand/finger movement? Hypothesis: Visual spatial working memory is more accurate after hand/finger movement.

Approach

To answer the questions and hypotheses in the last section, a human-subjects experiment was performed with 16 participants. The participants completed a set of 5 cognitive tasks per day on two different days, performing a hand/finger movement activity on one day, and a resting activity on the other; a counterbalanced design. The results allow the assessment of certain cognitive functions and the effect of the hand/finger movement on those cognitive functions. The 5 cognitive measures included are Reaction Time Test, Word Recall, Oddball, Go/NoGo, and Dot Matrix [10], [16]. The

Reaction Time Test required the participant to repeatedly react to simple stimuli, allowing for the precise measurement of reaction time. Word Recall required the participants to memorize a list of 30 words in 60 seconds, and then asked to recall as many words as possible 210 seconds after the Oddball activity, thus providing a measure of episodic memory. Oddball required a participant to react to one of three different stimuli, measuring the cognitive function of categorization. Go/NoGo is similar to Oddball but requires the participant to ignore one stimulus. The ignored stimuli does not show up often (20% of the time), thus, the participant develops a habit which must be overcome when the stimuli does appear, effectively measuring impulse control. Dot Matrix is a series of two different types of problems that are interwoven with each other, providing a measure for visual spatial (VS) working memory.

Assumptions/Limitations

Participant demographics were a limitation of the research as the demographics mirrored that of military operators. This sample consisted of 16 participants (14 male) ages 28-38 (mean 27), all holding bachelor's degrees in science or engineering fields, and 14 having or currently working on graduate degrees. Given these similarities, the findings may be limited and not generalizable.

Two assumptions were made for the hand/finger warm-up activity. First, the activity was purposely designed to have very few commonalities with the cognitive measurement activities. While some commonalities did remain, such as all activities being performed on a computer, the warm-up was assumed to be different enough from the

measurement activities as to not be considered a related warm-up task. Achieving separation between warm-up and task is pivotal to the ability to answer the main thesis question. A related warm-up activity would introduce other improvement factors beyond the mere movement of muscles such as the brain thinking about the tasks ahead and the muscles practicing them.

The second assumption made was that the warm-up mimics the physical mouse and keyboard motion of real time strategy games, without the mental demand. The hand/finger warm-up activity could be accomplished without looking at the screen. It was designed this way as stimulating the participant's brain as video games do with visuals, goals, and rewards would introduce another possible cognitive confound, making it difficult to conclude that the any cognitive improvement seen was due to small-muscle movement alone.

Key Findings

The results show an up to eight millisecond (ms) on average increase in reaction time to the stimulus when participants perform the treatment activity prior. This mean increase in reaction time was significantly faster ($p < 0.001$ Reaction Time Test, $p < 0.001$ Oddball) than the same participants' mean reaction time after performing the control activity. The size of the effect as measured by Cliff's Delta was small in each measure, but present. Results also show that participants in the treatment condition achieved a non-significantly ($p = 0.0736$) higher score in the Word Recall activity than in the control condition for one of four scoring approaches. Warmup-related improvement

on Oddball categorization and Word Recall episodic memory measures is low enough to recommend further research. In addition, it was found that participants had varying intensities of effort in the warm-up activity. Results from the analysis of these different levels of intensity show that a more strenuous warm-up produces larger cognitive effects in reaction time with a significantly positive correlation coefficient of 0.5397. It was also shown that the effects of the warm-up activity wear off over time across all participants. When controlled for demographics such as age or years playing video games, the effect of the warm-up was statistically different and diminished at different rates. Participants who were older, had less video gaming experience, and started playing video games later than age 13 saw larger effects from the small-muscle warm-up. Participants who exercised more than 3 times a week also saw a greater effect from the warm-up than those who exercised less. These findings offer statistical and substantial evidence to the scientific community that hand and finger movement, unrelated to the task to be performed, does have an effect on reaction time, episodic memory, and categorization.

Thesis Structure

The remainder of this thesis is divided into four chapters. Chapter 2 provides a discussion of the literature that was reviewed. The literature review will focus on relevant studies and the methods and test designs used in order to determine if small-muscle movement affects cognitive performance. Chapter 3 discusses the main research questions and hypotheses of the experiment and then describes in detail the methods used and how analysis will be performed. Chapter 4 reports the results of the analysis

performed and provides a discussion of those results. Chapter 5 provides a conclusion reiterating the main processes used and discovered results. Chapter 5 also provides a lessons learned and future work section which discusses possible limitations to this thesis's findings. Appendices A and B provide a look at individual reaction time results. Appendix C provides the lists of words used in the Word Recall activity. Appendix D contains the surveys used to collect demographic information. Appendix E provides pseudo code for a time scoring algorithm used for Word Recall analysis. Appendix F provides the IRB approval letter for the use of human volunteers in this research.

II. Literature Review

Chapter Overview

Many researchers have studied cognitive performance. This chapter highlights those that relate most to the research being conducted in this thesis. The chapter will begin with a discussion of the gap that is left by the research conducted thus far followed by the cognitive measurement activities that have been used in these studies and the associated cognitive functions being measured. Following this will be an overview of the common test design practices. Finally, there will be a discussion on areas of uncertainty within these studies.

Research Gap

The focus of this thesis work is hand and finger movement done prior to an unrelated task and associated cognitive effects. This specific focus has yet to be studied. A few areas related to this have been studied which include exercise and its effect on cognitive performance, gum chewing and its effect of cognitive functions, and surgical warm-up and its effect on surgical performance. This section explores each of those related research efforts.

First, Chang et al. performed an analysis of 79 studies on small bouts of large motor exercise and the associated cognitive effects. In this study, the combined results of all 79 studies showed that there was a small statistically significant increase in cognitive performance following a bout of exercise [2]. Hogervorst et al. explored cognitive

performance after strenuous physical exercise and also found significant increases in performance [17]. Studies in large muscle movement and associated effects have been going on since the 1980's if not earlier [17]. While these studies do deal only in large muscles and not smaller muscles like hands and fingers, they do provide strong evidence that muscular exercise can positively affect cognitive performance. Whether that effect is present in smaller muscular exercise remained largely unstudied until research in gum chewing began, which is discussed next.

Gum chewing and its effect on cognitive functions may be among the most relevant studies to this thesis's focus, because gum chewing involves small-muscles and is unrelated to most cognitive tests used. Studies by Allen et al., Onyper et al., and Sakamoto et al. show chewing gum can have various positive cognitive effects such as improving immediate and episodic memory, visual spatial working memory, vigilance, and participant reaction time; refer to Table 1 [15], [18], [19]. However, gum chewing studies suffer from a potential for noise. Noise may be introduced by the various sensations the mouth experiences from different types of gum. One study showed that flavored gum which was not chewed but held in the mouth affected cognitive performance [15]. Although such noise was a big part of gum chewing studies, the activities used in this thesis experiment to measure cognitive functions were adapted from gum chewing studies because they remain the most recent and relevant to this thesis's focus.

The last area of relevant material is the recent studies on warm-up in the surgical

field. Lendvay et al. speak to the fact that athletes have been warming up for centuries...

“yet, surgery does not involve a prescribed warm-up or pre-surgical rehearsal, although it is a high-stakes profession drawing on intense psychomotor and cognitive efforts” [6].

Many variations of small-muscle warm-up studies have been done in the surgical field ranging in its relation to the actual task, including: cell-phone app warm-up [3], generic video games [20], custom-made laparoscopic representative video games [14], and highly representative virtual reality machines [6]. The cell-phone app warm-up may be the most unrelated warm-up when compared with the actual task, next to gum-chewing studies. The surgical studies offer a variety of techniques and methods for measuring task performance, however cognitive effects are not always a primary focus point.

Aside from gum chewing and surgical studies, the literature search did not produce any other papers relevant to the thesis topic. The large muscle exercise studies provide promise that similar effects can be found with hand and finger movement. Surgical studies provided insight into the combination of small-muscle movement in relation to the task being performed and its effect on performance. Gum chewing studies provide many cognitive measurement activities that were leveraged for this thesis experiment and will be further discussed in the next section.

Cognitive Measurement Activities and Associated Cognitive Functions

Cognitive measurement activities consist of tasks whereby individuals participate

within the constructs of an experiment. The purpose of these tasks within any given experiment considered here is to measure one or more cognitive functions. Cognitive functions include but are not limited to: reaction time, visual spatial (VS) working memory, episodic memory, attention, vigilance [21]. This section discusses the most relevant cognitive measurement activities conducted within the literature review and what the results were for the associated cognitive function.

Full and Distracted Attention Measurement Activity (Episodic Memory)

Episodic memory is a cognitive function where an individual is able to remember a certain event at a certain time [21]. Onyper et al., in their study on gum-chewing and cognitive effects, measured episodic memory of participants using an attention word test [15]. This activity was executed in two different versions. The first was a full attention activity, where participants were allowed to devote their full attention to the words while being displayed. The second was a distracted attention activity, where participants had to complete a task while viewing the words. In their particular implementation of the test, participants were shown 30 words displayed one at a time for 1.5 seconds with quarter second breaks in between. The words shown were 6-7 letter nouns as recommended in previous research. Participants were told to remember these words. In the distracted version of this activity, participants were required to press different keys to the beat of a metronome while viewing the words. To ensure visual spatial working memory was not being tested, participants completed another activity directly after viewing the 30 words. The other activity lasted 3 minutes, at which point participants were told to recall as

many words as they could remember within 2 minutes. The results of this activity showed a significant increase in episodic memory as a result of gum chewing for both the full and divided attention activities [15].

Dot Matrix Activity (Visual spatial working memory)

The cognitive function of visual spatial (VS) working memory tests the participant's ability to store information for a short period of time, using that information to accomplish a particular task. Onyper et al. uses a dot-matrix with grid task to

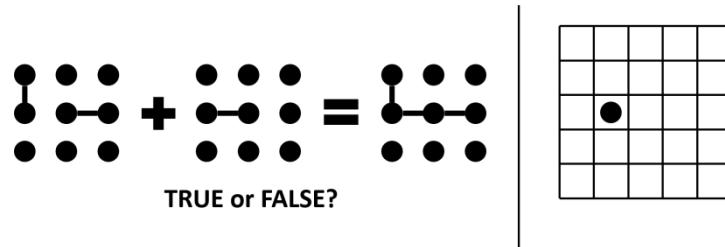


Figure 1: Dot-Matrix with grid

The dot-matrix is an equation made from dots in which a participant views and determines if it is TRUE or FALSE. To the side or shown after selection is a 5x5 Grid with a dot in it for later recall.

accomplish this as shown in Figure 1 [15]. In this particular set up, participants worked through 15 individual sets. The first 5 sets contained 2 dot-matrix equations each, with every equation being followed by a 5 x 5 grid with a dot placed randomly in one of the 25 grid spaces. At the end of each complete set, participants identify where on the grid the random dots were shown. This process continues for the remainder of the sets. The next phase of 5 sets uses 3 dot-matrix equations with grid and the last 5 uses 4 [15]. Visual spatial working memory objective is achieved by the participants being required to remember the locations of the random dots shown on each grid and report all locations at

the end of each set. The dot-matrix equation serves to add a level of complexity to the grid memorization problem by making the participant perform basic math. The results showed a significant increase over the control group in performance in VS working memory.

Symbol Digit Modalities Activity (Reaction time)

The last cognitive function discussed here which is used by Onyper et al. addresses reaction time of participants using a Symbol Digit Modalities test [15]. Here participants are trained on symbols and their associated number key 1-9. During the test participants are shown a symbol which appears at some random rate. Participants must map in their mind the symbol to the correct associated key and press the key as quickly as possible [15]. This test provides a more complex method to acquire reaction time, requiring the participant to process information before responding. The test also allows for manipulation to fit desired parameters. For instance, eye movement could be kept to a minimum by only showing the symbol in one central location. Key presses are limited as the participant is only hitting the associated key at the desired rate of display. The results show a significant increase over the control group in performance for reaction time [15].

Repeated Digits Activity (Reaction time)

Allen et al. uses a Repeated Digits test to capture vigilance [18]. This is a simple activity where participants keep their eyes focused on the center of a screen watching 3 digit numbers being display at a rate of 100 per minute. At times certain 3 digit numbers were repeated on the screen. Participants identify the repeat by hitting a key [18]. This

activity achieves the activation of the cognitive function by requiring the participant to watch the numbers very closely as they look for repeats. The results show a significant increase in performance for reaction time [18].

Oddball (Categorization and Reaction time)

Sakamoto et al. studied the effect of chewing on categorization in a time based activity [19]. The name of the test is called oddball and can be done in several different fashions. At its base, oddball targets one of the five senses with two different stimuli. The stimuli are presented in a random sequence, with the second stimuli occurring at a significantly lower rate than the other. When the second stimuli occurs, the participant must identify it [19]. This method can be implemented on any of the five senses. For instance, two audio sounds could be used, wherein the participant identifies when the second sound occurs. A similar activity was used by Cooper et al. for measuring the effect on reaction time [16]. In this study, oddball showed a significant effect over the control group on reaction time.

Go/NoGo Measurement Activity (Inhibition or Impulse Control Accuracy)

Cooper et al. measured the cognitive demand the brain experiences while performing inhibition or impulse control between two different stimuli using a go/no-go activity [22]. This activity consists of the participant responding to stimuli in a binary fashion. A series of stimuli are displayed and the participant is told to respond to each stimuli by pressing a certain key, these are go stimuli. For another stimulus, the participants were told not to press any key, this is the no-go stimuli. No-go stimuli occur

with less frequency than the go stimuli [22]. As the participant controls the urge to hit a key for the no-go stimuli, the participant expresses impulse control.

Switch Measurement Activity (Attention Accuracy)

Cooper et al. also measured the cognitive demand from different stimuli the brain experiences during an attention activity [22]. Participants work through a series of trials. Each trial is the same in how it is delivered, but the correct way to respond is altered. Between each trial a stimulus indicates whether to switch the manner in which to respond, or to continue to respond the same as before. This activity took the stimulus that was used as a distractor in the go/no go activity and gave it a specific meaning base on its orientation, which meant the participants had to pay attention to it [22].

Test Designs

In addition to cognitive measurement activities, the overarching design in which these activities were used was also looked at in the literature review. The most common test design used was a randomized controlled test [2], [3], [5], [6], [14], [15], [18], [19], [23]. In this design participants were randomly divided into two groups, a control and treatment group. The control usually entailed the participant performing some sort of normal or relaxing activity while the treatment was an activity such as gum chewing. While most studies follow this standard test design, the exact implementation varies among them and some are worth noting.

The first of these is found in a paper by Lendvay et al., which discusses two

concerns which involved the cognitive effect of imaginary warm-up and the learning effect [6]. In their study, participants in the control group were not able to start the task immediately. Instead they were told to read a leisure magazine for 10 minutes before the task. This aided to minimize the chances of the surgeons thinking about the task ahead. The authors addressed concerns about a learning effect by training each surgeon on the virtual reality machine to a certain proficiency benchmark [6]. This is the idea that it takes everyone different amounts of time to learn how to ride a bike. This idea of a benchmark serves to remove or reduce any possible learning effect.

The next method worth noting is from a paper by Onyper et al. which took a unique step back from the gum chewing studies and addressed the question:

“if chewing gum is indeed associated with increases in arousal and a corresponding shift in cognitive function, it is unclear why many studies find little to no performance advantages of chewing” ([15] p1).

The authors found too much variation in the results coming out of gum-chewing studies and aimed to boil it down to the basic variables that produce an effect and control them. Onyper et al. took what they believed to be these key dependent variables from other studies and built their test around those. Five cognitive tests were selected. Knowing from prior studies that any warm-up effect tends to die out around 15-20 minutes, the tests were fit within that window. To account for bias towards individual tests, the series of 5 cognitive measurement activities were strategically randomized. Results from this basic approach proved useful as they are able to find significant effects and show that gum chewing is only beneficial if done before, not during the activity [15].

The last method is a reoccurring theme that warm-up benefits the individuals with more task experience. Jalink et al. noted a surgical study which showed that only senior level residents, not junior residents, performed a suturing task significantly better after a warm-up [14]. Lendvay et al. discusses an unexpected finding that the MIS (minimally invasive surgery) experience of surgeons was a factor as warm-up appeared to benefit the more experienced surgeons [6]. Rosser et al. discusses a large-muscle warm-up finding where untrained individuals do not obtain the same physiologic and metabolic benefits from short periods of warm-up as trained individuals do, suggesting trained surgeons may benefit more than untrained [24]. Although this thesis study will not involve tests that any participant will have seen before and thus be more experienced, it is known that certain individuals may possess past developed skills which will lead to higher scores on cognitive tests. The Reaction Time Test activity requires participants to press the 'space' key as soon as an 'X' appears in the middle of the screen. This quick response like action is similar to experiences in games such as first person shooters where individuals quickly respond, though the aiming and movement is removed. The warm-up activity designed for this study also mimics the actions of a fast paced mouse and keyboard video game. Participant gaming experience was tracked to determine if the warm-up activity benefited a particular group. In another study by Rosser et al., video game skill was used to classify participants into 3 tertiles and found that the top tertile had 47% less errors when compared to the bottom tertile [20].

Results

Although small-muscle hand and finger movement and cognitive effects have not been directly studied, many of the results found in the few related topics provide substantial evidence that a correlation does exist. Table 1 lists these findings and their references as well as instances where no correlation was found. The following is a highlight of the results from the literature survey.

Table 1: Cognitive Effects and Treatments Used - The table displays Cognitive Effects found in the literature survey and the treatments used to study them. An ‘↑’ indicates there was a significant effect on the particular cognitive function, given the treatment. A ‘≈’ indicates no significant effect found.

Cognitive Effect	Treatments			
	Chewing Gum Before Only	Chewing Gum During	Related Task Warm-Up	Chewing In between
Episodic Memory	↑[10]	≈[10]		
VS Working memory	↑[10]	≈[10]		
Executive Function	≈[10]	≈[10]		
Reaction Time	≈[10]	≈[10], ↑[25]		↑[19]
Vigilance		≈ [25]		
Attention		≈ [25]	↑[5]	
Full Attention	↑[10]	≈[10]		
Divided Attention	↑[10]	≈[10]	↑[5]	
Decision-Making	↑[10]	≈[10]		

Onyper et al. found an increase in cognitive function in the areas of episodic memory, visual spatial working memory, and perceptual speed of processing (full attention, divided attention, and decision-making) [10]. The results were only true when chewing gum occurred before the test and not during. Results were found using statistical *t*-tests and the associated *F*-measure and *p*-value. However, the authors were unable to

show a statistically significant cognitive effect for verbal fluency. Onyper et al. also found that any cognitive benefits from gum chewing wear off within 15-20 minute of the activity [10]. The results from this study help to clear up what factors may cause increases in cognitive performance and what factors or methods may be causing them to decline or remain stationary [10].

In the paper by Allen et al. cognitive vigilance and accuracy tests were used, as well as electroencephalography (EEG) [25]. Through the use of a mixed ANOVA, reaction time (a characteristic of vigilance) was found to be significantly shortened over the control group. EEG sensor positions T3 and F7 beta activity rates were shown to have increased significantly for post-chewing, showing that frontal and temporal area activity in the central nervous system has a relationship to a person's vigilance [25].

A custom-made Wii Game developed for preoperative warm-up showed positive results in the paper by Jalink et al. [14]. Using *t*-test analysis, test scores for the peg transfer task (consist of lifting rubber rings off of pegs and transferring them) were found to have a significant increase in task performance for the treatment group verses the control group. The same results were shown for a test group for the other task called cobra drill. However the performance was a measure of seconds taken to complete each task and did not consist of a specific cognitive function [14].

Lendvay et al., who studied participants using a laparoscopic virtual reality machine as a warm-up tool found statistically significant increases for several factors [6]. The study was broken down into 5 separate factors which included task time. Two of the

factors including task time were shown to be significantly better for the treatment group. This test also only focused on task performance with none of the 5 factors consisting of a specific cognitive function. Not all 5 factors saw improvement. Fisher's exact-test and *t*-tests were used in this study as well as a repeated measures ANOVA model [6].

Areas of Uncertainty

As this research focuses in on an unexplored area, many unknowns are expected. However, even the well-researched gum chewing and surgical studies still have areas of uncertainty. This section discusses two of those areas.

The first area of uncertainty comes from a paper by Onyper et al., where the theory of dual-mechanism is discussed briefly as a possible explanation as to why many gum-chewing studies see no positive cognitive advantages [15]. It is proposed that the detrimental effects of chewing gum while performing tasks counteracts, or cancels out the positive effects of chewing gum as a warm-up. This is the primary reason why Onyper et al. performed the same study twice, varying only the treatment group by having one treatment group chew gum before the cognitive tests, while the other treatment group chewed during the tests. Only the study with the treatment group that did not chew gum during the tests had any significant cognitive effects as shown in Table 1 [10]. However, in the study by Sakamoto et al., the warm-up was repeated multiple times in-between each test [19]. Sakamoto et al. found a significant positive cognitive effect on reaction time in their study while Onyper et al. found no effect for participants chewing during tests [15], [19]. The area of uncertainty becomes the distance between continuing

the small-muscle motion during tests and performing the motion in-between the tests. If small-muscle movement during tests is detrimental to cognitive functions as Onyper et al. theorize, then where is the line drawn [15]?

The second area of uncertainty, which is applicable for most any test is that which may cause a placebo like effect. The gum chewing studies may have a higher inclination towards this than others. Like other gum chewing studies, the paper by Allen et al. leaves room for such an effect to occur. Telling someone to chew a piece of gum while performing a task could have such an effect on an individual. The individual may believe gum or mints improve cognitive performance. Knowledge that peppermint may help with attention and other cognitive function has been around for a while and participants may have been pre-disposed to this knowledge [26]. If a participant thinks the gum will help them, then we have a possible placebo effect. A possible remedy to the study by Allen et al. would be that the control group, instead of not chewing anything, be given a hard mint before the experiment. This could trigger similar placebo effects if they are occurring but would not require any mouth motion.

Summary

The research community has conducted many studies on exercise/warm-up and cognitive effects. However, these do not tend to explore a small-muscle cognitive effects. Gum chewing and surgical studies dominate the field of study. These studies and the techniques used are relevant to future work with small-muscle and cognitive effects. The literature provided many experimental methods to measure different cognitive functions,

and in many cases were successful in showing effects.

III. Methodology

Chapter Overview

This chapter addresses the methodology behind the collection of data for a human cognitive function experiment. This research will provide further knowledge to the human machine teaming community by asking a specific question about small-muscle exercise and its effect on cognitive performance. If found that there is an effect, many possible operator systems can use this information to their advantage to improve operator performance. The chapter first discusses the research questions in regards to small-muscle exercise and cognitive effects. Cognitive functions should increase or decrease in performance when small-muscle exercise is performed prior. Exactly how prior small-muscle exercise is estimated to affect cognitive functions is discussed next followed by how those cognitive functions can be measured. The chapter will end with an overview of the tasks that will be used to measure the cognitive effects and the analysis process.

Research Questions

The purpose of this research is to determine if there is a cognitive effect from hand/finger movement done immediately prior to a non-related task. There are a wide range of cognitive functions which prior hand/finger warm-up may affect. In this research, cognitive functions measured are reaction time, categorization, inhibitory/impulse control, episodic memory, and visual spatial (VS) working memory. For each cognitive function one or more questions are asked. The following paragraphs

further describe our beliefs about cognitive function and the questions asked for this research. Table 2 provides summary of the research questions, and includes a tag for locating all discussions in regards to the tagged question.

Table 2: Research Questions – Column 1: Research questions. Column 2: Tag for locating discussions in regards to cognitive function and research question. Column 3: Cognitive function in question.

Questions:	Tag:	Cognitive Functions (CF):
Does hand/finger movement immediately prior to a non-related task improve CF?	T1 T2 T3 T4 T5	Reaction time Categorization Inhibitory/Impulse Episodic Memory Visual spatial working memory
Do CF improvement gains diminish over time?	T8	Reaction time
Is a learning effect present from Day 2 to Day 3 in CF?	T9	Reaction time
Does greater hand/finger movement during treatment activity increase the size of the effect?	T10	Reaction time

The cognitive state of an individual can affect their cognitive performance. Many studies have shown that activities can be performed to directly improve cognitive functions. Caution must be taken as there is much potential for inadvertently introducing confounding factors. Gum chewing introduces different confounding factors such as taste and tasteless chemicals, which could affect the results of the study. Despite this it is believed that if all confounding factors were eliminated, the activity of moving small-muscles, by itself, would improve cognitive performance. This research asks if hand/finger movement done immediately prior to a non-related task improves cognitive functions including reaction time, categorization, inhibitory/impulse control, episodic

memory, and VS working memory.

As discussed in Chapter 2, Onyper et al. found that the effect of the treatment wears off over time [10]. It is unknown if a similar diminishing effect would be present after hand/finger exercise. This research also asks if reaction time, categorization error count, and inhibitory/impulse control error count improvements diminish over time.

This research also looks for the possibility of a learning effect. Because the study takes place over a period of days, and the activities are repeated, a learning effect is likely. Though it can be reduced through a counter balanced test design, but that should not serve as a replacement for actually testing for a learning effect.

Lastly, this research looks at the intensity of the hand/finger exercise. Participants vary in their abilities with the mouse and keyboard and while participants are encouraged to achieve and maintain a standard rate, some participants will fall further away from the desired standard during the hand/finger warm-up activity. This research asks if those who are more active see a larger effect by the treatment.

Hypotheses

The results generated are expected to show a positive cognitive effect from hand/finger movement done immediately prior to a non-related task for five select cognitive functions. Therefore, each of the questions described in the prior section has an associated alternative hypothesis which in general states that an improvement in that cognitive function will be seen. Table 3 provides a list of these research hypotheses.

Table 3: Research Hypotheses – Column 1: Research alternative hypotheses. Column 2: Tag for locating discussions in regards to cognitive function and research hypotheses. Column 3: Cognitive function in hypothesis.

Alternate Hypothesis:	Tag:	Cognitive Functions (CF):
CF is faster after hand/finger movement.	T1 T6	Reaction time Episodic Memory
CF has fewer errors after hand/finger movement.	T2 T3	Categorization Inhibitory/Impulse
CF is more accurate after hand/finger movement.	T4 T5	Episodic Memory Visual spatial working memory
CF is more accurate and faster after hand/finger movement.	T7	Episodic memory
The effects of treatment activity on CF will diminish over time.	T8	Reaction time
A learning effect is present from Day 2 to Day 3 for CF.	T9	Reaction time
Greater hand/finger movement during treatment activity does not increase the size of the effect of treatment.	T10	Reaction time

Dependent Measures

(T1, T8, T9, T10) Reaction time: This is time it takes to respond to stimulus. The stimulus does not always have to be visual as is used in the experiment for this thesis. A stimulus can be an audible noise, a physical touch, a particular task, or a smell. The response likewise can be made in various ways, but should be able to be done quickly, like a key press, as to not add unnecessary time. The dependent variable is reaction time. Reaction time can be affected by both fatigue and learning effects. As with any cognitive

function, reaction time may be affected by the amount of muscle movement during an exercise treatment done before the task.

(T2, D2) Categorization: In Cooper et al. study, participants were asked to classify colors as they were presented in series [22]. This process is a form of categorization where stimuli must be categorized. The smaller the difference between the stimuli, the more difficult it will be to correctly classify them. An improvement in categorization would constitute a reduction in the number of the stimuli being misclassified. The dependent variable is the number of stimuli incorrectly categorized.

(T3) Inhibitory/Impulse Control. If we repeatedly perform a particular action in response to a stimulus, our brains make connections between the stimulus and the associated action such that we do not even need to think about it [27]. Once these habits are formed, it becomes difficult to prevent this impulse to respond when the stimulus is shown again. This act of preventing the action is known as inhibitory or Impulse Control. If an individual is told no longer to perform a habitual action when a particular stimulus is shown, a reduction in the number of habitual response errors would constitute an improvement in Inhibitory/Impulse Control. The dependent variables are the number of stimuli incorrectly controlled for.

(T4) Episodic Memory: As described by Tulving et al., episodic memory consists of temporal events that are received and stored in the brain [28]. These events can be physical, emotional, or any other experience. The main key is the storage of an event that occurred. Later, the participant can be asked to recall as many events as possible. It is

through the tallying of the number of events recalled that provides a measure for episodic memory. An increase in the number of events recalled would constitute an increase in episodic memory performance. Variations in attention can be achieved by requiring the participant to perform other actions during the time the events are presented. This is known as a divided attention test and was done by Onyper et al. in which participants pressed a key to the beat of a metronome while experiencing each event [10]. A full attention task would not include such distractions. The dependent variable is the number of correct responses.

(T6, T7) Whenever a participant is asked to recall events, there will be some delay between being asked and the actual recall of each event. This is referred to as episodic memory recall speed. It is physically impossible to recall multiple items at one time, or enter them for logging purposes simultaneously. This creates different times of entry for every event recalled. A distribution of times in which words are entered is thus produced. The recall speed is the sum of these times of entries. An improvement in episodic memory recall speed will consist of improved distribution being more shifting further right than the original distribution, thus a lower total recall speed.

(T5) Visual spatial (VS) working memory: This is often thought of as a limited information bank which takes temporary input that can be manipulated for purposes of solving a problem [29]. Because the storage is limited, the individual must respond within a reasonable time so that information is not lost. However, some amount of time should pass between when a stimulus is experienced and when the individual is to recall

the stimulus so that storage does occur. The insertion of another task during this delay in response can help force the brain to temporarily store and manipulate the given and new information as to correctly respond to both problems. An increase in the number of correct responses constitutes an improvement in visual spatial working memory. The dependent variables are the number of correct responses.

Cognitive Measurement Activities

For all dependent measures and variables listed in the previous section, cognitive measurement activities were developed for purposes of quantifying the dependent variables. A Reaction Time test activity is used to capture reaction time (T1, T8). Categorization is captured through an activity known as Oddball (T2). Inhibitory/Impulse Control is measured through an alternate version of Oddball called Go/NoGo (T3). Episodic memory is measured through a simple Word Recall activity that includes a pause between viewing and recalling the words (T4). Visual spatial working memory is measured through an activity called Dot Matrix. These cognitive measurement activities are outlined in detail in the following paragraphs.

Measure 1 - Reaction Time Test (Processing Time: T1, T8)

This is a reaction time focused measure in which the participant is first shown a blank white box in the center of the screen. After a delay time which has been selected at random, a black 'X' stimulus appears in the white box as shown in Figure 2. The delay time is randomly chosen from a list of three delay times: 0.5 seconds, 1 second, and 1.5

seconds. The participant reacts to the presence of the 'X' by pressing the 'space' bar as soon as they notice the 'X'. The 'X' will remain in the box until the participant presses the 'space' bar. The time from which the 'X' appears to the time the participant presses the 'space' bar is recorded. The process then repeats itself for a total of 240 seconds, with the participant experiencing approximately 145 instances. Because participants vary in reaction time, a different number of instances are experienced between treatment and control groups and between participants, though the order is the same. Thus when comparing two groups, the minimum number of stimulus experienced from the two groups is used for both producing an equal number of same order instances for analysis.



Figure 2: Reaction Time Test Example - Participants are shown a white blank box. At random times an 'X' appears in the box at which participants react by pressing the 'space' bar. The time taken to react is recorded.

Measure 2 - Oddball (Categorization T2, D2, G2 [22])

This is an activity used to measure categorization error. In this activity, three different colored Gabor patches (red, black, blue) are displayed one at a time either center right or center left (shown in Figure 3), at specified probabilities. Red Gabor patches show up 40% of the time, blue Gabor patches for 40%, and black Gabor patches 20%.

For each colored Gabor patch, the number of lines and orientation may vary. The Gabor patches' colors, line variations, and orientation variation is shown in Figure 4. A Gabor patch is displayed for a total of 0.1 seconds with a 1 second delay between patches. In the Oddball version, blue and black patches are ignored while the participant presses the 'right control' key when a red patch is displayed. If a red Gabor patch is displayed and the 'right control' key is not pressed, a feedback message is displayed that reads 'Too Late,' indicating to the participant that they should have reacted. If the 'right control' key is pressed when a blue or black patch is displayed, or any other key is pressed besides the 'right control' key when a red Gabor patch is displayed, a feedback message is displayed that reads 'Incorrect,' indicating to the participant that the wrong reaction was taken. Each time the 'right control' key is pressed the result is recorded. The Oddball activity runs for 210 seconds, 30 seconds less than Reaction Time Test or Go/NoGo in order to reduce the time between word memorization and recall. This reduced time was chosen to reflect similar times used by Onyper et al., and because of the inherent difficulty of the word recall task.



Figure 3: OddBall and Go/NoGo Display - Shown is a red Gabor patch in center right position. Gabor patches show in either center, left, or center right.



Figure 4: Six Gabor Patch Variations - These are the 6 Gabor patches a participant will see in the Oddball and Go/NoGo activities. Red and blue Gabor patches equally make up 80% of the patches displayed. Black Gabor patches make up the remaining 20%. The participant responds to the color of the Gabor patch, not the orientation or number of lines.

Measure 3 – Go/NoGo (Inhibitory Control T3 [22])

The Go/NoGo activity visually looks and operates the same as the Oddball activity, with a few differences. The Go/NoGo activity requires the participant to press the ‘left control’ key when a blue Gabor patch is displayed in addition to pressing the ‘right control’ key when a red Gabor patch is displayed. This means for approximately 80% of the time the participant is reacting to the stimuli, creating a habitual response. The participant receives the same feedback as in Oddball with the addition of a message stating “Too Late” when a blue Gabor patch is displayed and no key is pressed. The feedback message “Incorrect” is displayed also if any other key besides ‘left control’ is pressed when a blue Gabor patch is displayed. Each time the ‘right control’ or ‘left control’ keys are pressed the result is recorded. The Go/NoGo activity runs for 240 seconds.

Measure 4 – Word Recall (Episodic memory T4 [10])

The Word Recall activity is used to measure episodic memory. A series of 30 words is displayed one at a time for 52.5 seconds. A gray bar is below the word indicates

how many words are remaining as shown in Figure 5. Each word is displayed for 1.5 seconds followed directly by a 0.25 second break before the next word. The words are nouns and verbs ranging in length from 6-7 characters. The words came from the MRC Psycholinguistic Database, and have concreteness, familiarity, and image ability ratings of 400-650 [15]. There are 12 lists of 30 words with some words repeating between different lists for a total of 204 unique words (see Appendix C). Participants see one list of 30 words in treatment condition, and another in the control condition. The order in which two lists are used for one participant is swapped for another participant to mitigate recall effects caused by lists of words that are easier to recall. Once all words are displayed, the Oddball activity is run which serves the purpose of inserting time between memorization and recall of the words. Once the Oddball activity is complete the participant is asked to recall as many words as they can within 120 seconds by typing them in one at a time. All participant input is recorded.



Figure 5: Word Recall Display - Participants see one word for 1.75 seconds followed directly by a 0.25 second inter-stimulus interval. Gray boxes below the word indicate then number of remaining words.

Measure 5 – Dot Matrix (Visual spatial (VS) working memory [15])

The dot-matrix activity consists of dot matrix equations and dot grids. A dot matrix equation is an addition problem in which dots are connected by lines in two

separate 9 dot matrices. A third 9 dot matrix represents the possible answer to the addition of the other two. A participant must decide if the lines in the first two matrices are correctly added in the third. Figure 6 shows such an equation, which in this case evaluates to true.

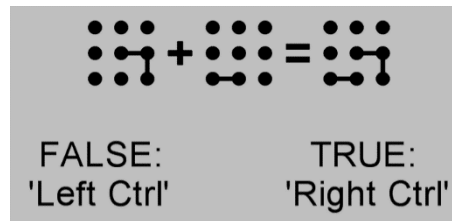


Figure 6: Dot Matrix Equation – What a participant sees. If the lines in the grid right of the equals sign match the addition of the lines in the two other grids, the equation is TRUE. If it does not, it is FALSE.

The dot matrix equation will show for 10 seconds in which time the participant must decide and indicate if the matrix is true or false. The participant does this by hitting the ‘left control’ key for false and the ‘right control’ key for true. The participant’s response is recorded. If the participant does not respond before the 10 seconds has expired, that information is recorded as an incorrect answer. Once the participant has responded or the 10 second timer runs out, a dot grid is displayed for 1.5 seconds. As shown in Figure 7, a dot grid is a white 5x5 grid with a black dot located randomly in one of the 25 positions.

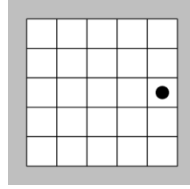
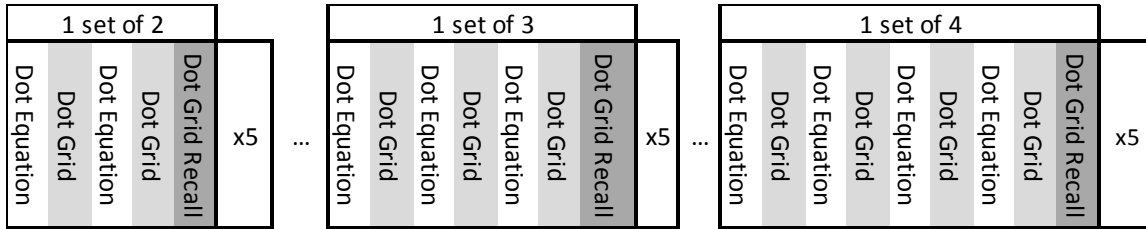


Figure 7: Dot Grid Display – A white 5x5 grid with a random black dot in one of the 25 positions.

The participant is instructed to memorize the location of this dot within 1.5 seconds. After this the participant is shown a black distractor mask for 1 second and then another dot matrix equation followed by another dot grid with a dot in another random location. This series of dot matrices and dot grids completes 1 set of 2, meaning a set with 2 dot matrix equations and 2 dot grids. The sizes of the sets used in this activity are 2, 3, and 4. At the end of a single set, the participant is shown a blank dot grid and is instructed to indicate all locations of dots seen during that set. Once the participant indicates the locations by clicking with the mouse in those locations, the participant clicks submit and the next set is started. This series of sets is shown in Table 4. The activity has a total of 15 sets comprised of 5 sets of 2, 5 sets of 3, and 5 sets of 4 for a total of 45 dot matrix equations and 45 dot grid locations, always shown in the same order. All responses are recorded.

Table 4: Dot Matrix Series of Sets - From left to right, participants experience each set 5 times before continuing on to the next larger set.



Treatment Conditions

Treatment - Small Motor Movement (warm-up):

Unlike cognitive performance experiments with large muscle exercise as the treatment [13], the participants in this experiment must move their hands/fingers at a reasonable rate (though variations occur) that would mimic either video game play or military operator controls such as piloting an unmanned aerial vehicle. This treatment was designed to do that without being mentally engaging. Figure 8 displays what participants saw during this activity. It consists of a minimum box indicator (1), key indicator boxes (2), and an activity gauge bar (3). The minimum box indicator (1) is a white bracket box used to indicate the minimum size box that can be drawn and still count. The brackets turn green if a correct box is drawn. The key indicator boxes (2) are squares that display the letter 'A' in the left box, and 'D' in the right box, turning green while displaying the letter. In this example, once 'A' is pressed, the box changes to white and the other box displays 'D'. The activity gauge bar (3) is filled with dark grey and has a white border. The bar contains a marker and a desired rate region. The marker is a small

darker grey rectangle while the desired rate region is a larger green rectangle. The marker moves left and right within the boundaries of the bar to indicate the current activity rate.

A participant's goal is to keep the marker in the desired rate region.

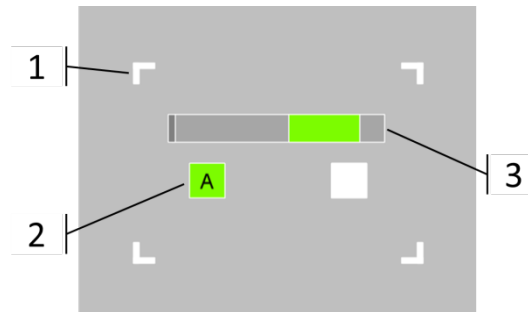


Figure 8: Treatment Activity

For 80 seconds, participants attempt to keep an activity measurement gauge (3) maker in the desired rate region. Participants use their left hand to alternately hit the 'a' and 'd' keys while at the same time repeatedly clicking and holding the left mouse button and dragging a white bordered box around minimum box indicator brackets. If the same letter is pressed twice or more in a row, it does not count towards the participants rate and the key indicator boxes (2) remain unchanged. Once the correct key is pressed, the current green indicator box turns white and the other key indicator box is turned green displaying the correct key to press. If the participant draws a box that has any part of its border inside the minimum box size indicator (1), or it does not encompass the minimum box indicator (1), it does not count toward the participant's rate. It is the combination of the two activities done correctly and at the correct rate that keeps the marker in the desired rate region. The rate is calculated based on the number of correct key presses and box draws and then scaled such that the majority of value fall between zero and one.

Each observation used to calculate the rate looks backward and determines how many key presses and how many mouse events occurred in the last two seconds. Referring to Equation 1, the total number of key presses plus the total number of mouse events plus the product of the two are added together, but not before each of them are multiplied by individual weights. The weights are used to give equal value to the individual mouse and keyboard events (alpha and beta) while ensuring the combination of the two events is rewarded higher (gamma) as it is desired that participants move both hands simultaneously. This generates a score which is then divided by the maximum likely score possible, generating the rate for that observation. The larger difference between alpha, beta, and gamma do not indicate more weight, but are needed as the manner in which mouse events and keyboard events are initially scored varies. Thus, mouse events require a beta value of 240 to be equivalent in weight to key presses with an alpha value of 60.

Equation 1: Activity Rate Calculation

$$Rate = \frac{\alpha * KeyPresses + \beta * MouseEvents + \gamma * KeyPresses * MouseEvents}{Maximum\ Likely\ Score}$$

Where $\alpha = 60, \beta = 240, \gamma = 30$

The desired rate region which is indicated by the green rectangle in Figure 8 is a rate between 0.65 and 0.85. While the participant’s actual rate is not shown to them, the

marker represented by the small dark grey rectangle in Figure 8 moves up and down the status bar to help the participant gauge their rate.

Keeping the participants within the desired rate region helped to move towards a standardized rate (though variations occurred), ensuring all participants receive similar small-muscle exercise. However, five participants could not adequately stay within this region. In an effort to not discourage these participants, a rate multiplier handicap was calculated and applied to their rate before day 2 and day 3. This was conducted by visual inspection of each participant's treatment activity rate from the training day. If the participant was unable to consistently stay in the desired rate of 0.65 to 0.85, the test administer assigned a unique multiplier. The rate handicap allowed the participants to keep the indicator within the desired green region at their best rate but did not change the rate that was recorded. Participants were unaware of this rate multiplier as knowledge of the handicap could lead to a self-devaluation cognitive confound which might increase stress.

Control - Inactivity

The control condition's purpose was to keep the participants from performing any kind of mental or physical activity. In this condition, participants stared at a small oval in the center of their screen for 80 seconds (same as treatment duration). The oval changes color to help keep their attention but not stimulate their brains. No other action is required by the participant at this time.

Hidden Factors

Commonly seen but difficult to account for factors include fatigue and learning-effects. This experiment was designed to provide enough data points such that adequate analysis of cognitive effects could be performed. As such activities were a minimum of 210 seconds and testing phases ranged from 25-35 minutes among participants. Due to the length and repetitiveness, fatigue was likely, but little could be done in test design to reduce it as high numbers of observations were needed, which equated to an approximate 30-minute experiment duration on both day 2 and day 3. Fatigue may be detectable through a comparison between the control and treatment groups for a particular cognitive measure of interest such as reaction time. If both groups have similar reductions in performance over time, it could be due to fatigue. At the same time, the presence of a learning effect could also be present. A learning effect could show itself over the duration of a single activity or between day 2 and day 3 of the experiment. Any possible learning effect was minimized through experimental design by training each participant until they could adequately perform each task such that they were not surprised or needing to learn anything on day 2 or day 3.

Experiment Sequence

This experiment was conducted under an Air Force Research Laboratory's Wright Site Institutional Review Board (IRB)-approved protocol (approval number: FWR20160127H, date: 25 Aug 2016. See Appendix F: Approval Lette). The experiment was conducted over three days for each participant. The days were not necessarily

consecutive with the majority of participants completing all three days within a seven-day span, and no participant exceeded a 14-day span. For days 2 and 3 the participants were scheduled for the same time of day as to not introduce any confounding factors due to differences in time of day. On day 1, participants were trained on each of the five measurement activities as well as the treatment and control activity. The training gave the participants the ability to become familiar and proficient with all activities. The proficiency of a participant was gauged by the test administer through observation. This was important as it reduces the chances of learning effects between day 2 and day 3. The actual experiment took place on days 2 and 3. Both day 2 and 3 received the same sequence of activities, except that the treatment and control were switched. The sequence consisted of each activity being preceded by 80 seconds of the treatment or control as shown in Figure 3.

Table 5: Experiment Sequence - Shows the sequences of activities for each day. The order of treatment and control was counterbalanced across day 2 and day 3.

Treatment (Day 2 or 3)	Control (Opposite Day)	Duration
Treatment	Control	80 seconds
Reaction Time Test	Reaction Time Test	240 Seconds
Treatment	Control	80 seconds
Word Recall (Display)	Word Recall (Display)	54 Seconds
Oddball	Oddball	210 Seconds
Word Recall (Input)	Word Recall (Input)	120 Seconds
Treatment	Control	80 Seconds
Go/NoGo	Go/NoGo	240 Seconds
Treatment	Control	80 Seconds
Dot Matrix	Dot Matrix	10 to 15 minutes

This process of repeating the treatment or control before each activity was adapted from a study by Sakamoto et al. and is has been shown to be the most effective

strategy for observing positive effects from a given treatment on cognitive performance [19]. Participants took a pre-experiment survey on both day 2 and 3 as well as a post experiment survey on day 3 (see Appendix D).

Assumptions

Literature in the area of small-muscle exercise is limited with regards to its effect on cognitive performance. Studies involving Gum chewing are the only source of information which involves both a small-muscle movement and associated cognitive performance [10], [19], [30], [31]. Because of the limited literature, common practices in small-muscle test designs were sparse and thus required several assumptions:

- One to two minutes of treatment is adequate for small-muscle exercise to produce cognitive effects if in fact cognitive effects are produced by small-muscle exercise. The length of exercise is an important factor as too little will not induce any cognitive improvement, and too much can cause fatigue which may negate any cognitive improvements.
- The treatment can be accomplished via muscle memory alone and does not require cognitive thought once muscle memory takes over participant's physical functions. A small-muscle warm-up treatment that is easy enough to allow the participant to perform without looking at the screen means the only requirement is the participant thinking about moving their hands. This is important in order to reduce other factors that could contribute to cognitive improvement.
- The cognitive measurement activities are adequate to detect any effect if it is

present. Effects of small muscular exercise are assumed to be less than that of large muscle exercise but equivalent to that of gum chewing. Thus the design of cognitive measurement activities that are sensitive to these smaller effects is pivotal to detecting their presence.

Data Collection

Data Collection is done via calls within PsychoPy code [32]. PsychoPy stores desired data in a comma separated value (CSV) format. Two files are generated: one to handle treatment data since it can grow large quickly, the second to store the 5 tasks. Data is stored in rows. The columns identify what is being stored. A 'task' row is used to identify which activity the participant is on.

Other participant demographic information was collected via survey at the end and start of the experiment and manually entered into comma-separated value (CSV) data files.

The PsychoPy package includes various methods for storing data. For this experiment data was collected and stored in an excel CSV format, then imported as individual data frames within 'R Studio' (Version 0.99.879, R x64 3.2.3) software.

Measure 1 - Reaction Time Test (Reaction time T1, T8 [33])

Data Cleaning: A byproduct of the Reaction Time Test measure design is that the faster a participant responds, the more stimuli will show up to react to, because the test

runs for exactly 240 seconds every time. This means one participant could see 144 stimuli, and the another 150 stimuli. To account for this, the lowest number of stimuli seen was used as a limit for all participant data. Thus any stimuli seen after 144 by any other participant was not used.

On occasion, a participant might accidentally react without being triggered by an actual stimulus – called a false response. This causes two unique data cleaning efforts. If the participant responded before a stimulus is presented, this causes an incomplete data entry. These data entries are removed. On the other hand, it is possible the participant accidentally reacts at the near millisecond time the stimulus displays, allowing for a data entry with a non-human capable reaction time. To handle these types of entries, a threshold was put in place. In a study by Jain et al. comparing visual reaction time to audible reaction time, the mean visual reaction time varied among the demographic groups within 120 participants from 219 ms to 256 ms [34]. With such variation in reaction times it was difficult to choose a particular threshold. However www.humanbenchmark.com has run a simple reaction time test with over 39 million data points. From charts provided we see that most individual's, approximately 95%, are above the 200 ms reaction time [35]. Thus this was chosen as the threshold and any data entry with reaction time less than this value was not considered. It still is physically possible to have a reaction time below 200 ms. To ensure none of the participants happen to be among the few that do have a reaction time below 200 ms, the mean and mode of all participants was inspected. From the analysis no participant had sufficient evidence to support having consistent reaction times below 200 ms, thus it was maintained as the

threshold for this experiment. If an entry was removed in the control or treatment dataset, the corresponding entry in the other data set (treatment or control) had to be removed also. For example, if data entry row '78' in the participant's control dataset was removed, data entry '78' in the participant's treatment dataset must also be removed to maintain the alignment and correct future analysis of data. 12 of the 4725 observations had reaction times below 200 ms, thus a total of 24 observations were removed. The maximum reaction time among all participants was 1614 ms. These extended reaction times were not considered lengthy enough to elicit removal.

Analysis Technique: The treatment group was first compared against the control group based on reaction time. This was done by performing a two-sample paired Wilcoxon Mann Whitney (Wilcox) test. The Wilcox test was chosen for all comparisons of means within this research. As opposed to a Welch's *t*-test, the Wilcox test does not make an assumption about the underlying distribution of the data and is thus an acceptable test for non-parametric distributions of data. The paired parameter is used because the participants are their own controls. This analysis provided information as to whether the treatment group did have a significantly faster reaction time than the control group. A similar test is also run on the day 2 and day 3 groups to test for learning effects. Further analysis is completed to determine the size of the effect using Cliff's Delta, pairwise subtraction, and a one-sample Wilcox test on the pairwise subtraction results.

The Cliff's Delta value is a useful tool for determining size of an effect, such as the treatment. Cliff's Delta and Cohen's D are both industry standards for measuring

effect size. However, Cohen's D assume the data follows a normal distribution. As with the Wilcoxon test, Cliff's Delta does not assume anything about the underlying distribution of the data and is thus an acceptable test for non-parametric distributions. Using this tool, a sliding window method is used to determine if the effect size diminishes over time. This is done by taking the first 0 thru 60 observations from both treatment and control data, and calculating a Cliff's Delta value for that window. Next the window slides to 1 thru 61 observations (a 59 observation overlap) and another Cliff's Delta value is calculated and stored. This sliding window continues until a new dataset of average time and Cliff's Delta values are created. The sliding window method used to determine effect size over time was also applied to different factors to gain more insight on demographic information. Four demographic factors were used (and split into two categories): Age at which the participant started playing video games (preteens vs teens or later), Age of participant (21-25 years vs 26-38 years), How many years the participant has played video games (0-15 years vs 16+ years), and number of days a week the participant exercises (1-3 days vs 4-6 days). From this dataset, plots are generated and used to visually inspect the effect size changes and Wilcoxon tests are used to determine if there is a significant difference in the effect between the front and second half of the activity.

Measure 2 & 3 - Oddball task T2 (Categorization), Go/NoGo T3 (Inhibitory Control) [22]

Data Cleaning: While in general, instructions were adequate in prepping participants for these tasks, one participant did not remember the instructions and thus the

first 75 seconds of observations were affected until the participant asked the experiment administrator for a reminder. For analysis, the 75 seconds of observations only needed to be removed for the individual that made the error. Only one other participant had such an issue and 15 seconds of observations were removed for that participant's data. To determine if such an issue is present, an inspection of each participant's response plot was conducted. Such a plot is shown in Figure 9, which displays the responses of the participant who asked for a reminder. The plot shows an unusually large number of incorrect responses in the beginning 75 seconds.

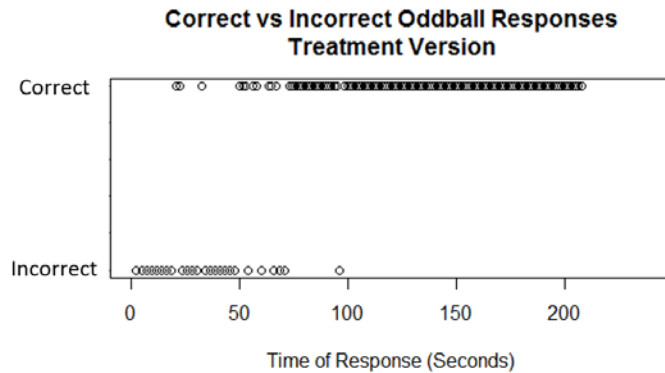


Figure 9: Time Series of Correct vs Incorrect Oddball Responses – The large quantity of incorrect responses in the beginning of this activity indicate the participant did not recall how to perform the experiment.

It is possible for participants to see a different number of stimuli as a result of the constant activity duration and mistakes generating feedback which adds time. For individual participant analysis, the minimum number of stimuli seen across the two versions (treatment and control) is used and thus varies between participants.

Analysis Technique: For each participant, a confusion matrix is generated such as

the one shown below in Table 6. The matrix shows the number of responses the participants made when they should have and should not have as well as when the participants ignored the stimulus when they should have and when they should not have. From this the number of errors will be calculated for each participant per experiment type.

Table 6: Example Go/NoGo Confusion matrix

	Should Press Red	Should Press Blue	Should Ignore
Responded Red	74	4	0
Responded Blue	2	66	0
Ignored	1	0	30

Once the number of errors have been determined and stored in a new data frame for all participants, categorized by experiment type, a paired Wilcox test will be conducted to determine if there is a significant difference between the treatment and control groups.

For the Go/NoGo activity only, inhibitory/impulse control error will be calculated by only considering the outcome of the black stimuli which is to be ignored. A paired Wilcox test will be conducted to determine if there is a significant difference between the treatment and control groups.

Measure 4 – Word Recall T4 (Episodic memory [10])

Data Cleaning: The fastest way to process this data was to manually inspect each input that was not a direct match and compare against the words shown. This is due to the

high number of misspellings or words that don't exactly match. For instance, if one participant entered the word 'aloo', if the word 'aloof' had been in the word list shown, a determination would have to be made as to whether to count that word. A script was developed so that each word inputted by the participant that was not 100% correct is displayed along with the list of words the participant saw for the purposes of scoring each word. To illustrate the scoring mechanism used, let us say the word seen was "monster." Direct matches were given a score of 3. If the word had a slight misspelling with 1 or 2 errors, such as "munster", it would be re-logged as a match and given a score of 2. If the word had more severe misspellings with 2 – 3 errors, such as "mounsters", but it was still very obviously the correct word, it would be re-logged as a match, and given a score of 1. If the word did not match, but did fit the category, such as "scary," it would be left logged as unmatched and given a score of 0. If it was an obvious different word and was not in the same category, such as "houses," it would be left logged as unmatched and be given a score of -1.

In addition to word score, the other item of interest in comparing the treatment and control groups was time of word entry. Upon inspection of the data, the treatment group was found to have inputted 38.3% of their total correct words within the first 10 seconds and 72.9% in the first 20 seconds verses the control group who inputted 38.5% of their total correct words within the first 10 seconds and 70.3% in the first 20 seconds.

Analysis Technique: For added insight into the different factors, the scoring and time of entry information were analyzed with four different approaches. A paired Wilcox

test was conducted for each approach to determine if there is a significant difference between the treatment and control group scores. For the first approach, the unscored words (no scoring information) were summed to determine how many direct matches a participant had in each group. The second approach used the scoring information to develop totals for each participant in each group. The third approach only considered the time of entry to determine a score for each group. Equation 2 is used to give higher scores for words entered earlier in the entry window. A tuning parameter m is chosen empirically and is used to alter the curve which affects how quickly words lose points for time of entry.

Equation 2

$$\text{Score} = e^{-\frac{x}{m}} * w$$

x equals time of word submission

Where m varies the value of a submission (set to 15)

w adds weight to the score (set to 6.6)

Observations of the data lead to the value of 15 chosen for m as it provides a sharp cutoff, giving entries within the first 20 seconds the highest scores. This was the desired effect as approximately 70% of correct entries were made within the first 20 seconds for both treatment and control groups, meaning participants were dumping all the words they had queued up from episodic memory. The remaining approximately 30% of entries were scattered throughout the next 100 seconds. Rewarding words in the remaining 100 seconds would not prove useful as the words are much more widely

scattered, thus a more forgiving scoring function might score a word at 30 seconds similar to one at 90 due to the spans needing to be covered. Another way to say it is that this time scoring function really only considers words entered in the first 30 seconds, and gives more value to words entered closer to start. A time score is given to each word. Thus each word a participant enters has a word score and a time score. w is a multiplier used to add more weight to the time score. Without 'w' a word entered at 1 second would receive a little less than 1 for a time score. As the sum of the word score and time score is calculated per participant per condition, the smaller time scores would be overshadowed by the word scores. Thus w is used to keep the two scores evenly weighted. This is only needed for approach 4 which uses w to create similar weights between the time of entry score and the word score. 6.6 was assigned to w as it provided similar weights between the two scoring mechanisms. It does not matter what w is set to for approach 3 as word score is not considered. The function with the discussed values, $m = 15$ and $w = 6.6$ is graphed in Figure 10. Using this graph, if a word was submitted at 20 seconds, the participant would receive an approximate score for that word of 1.5.

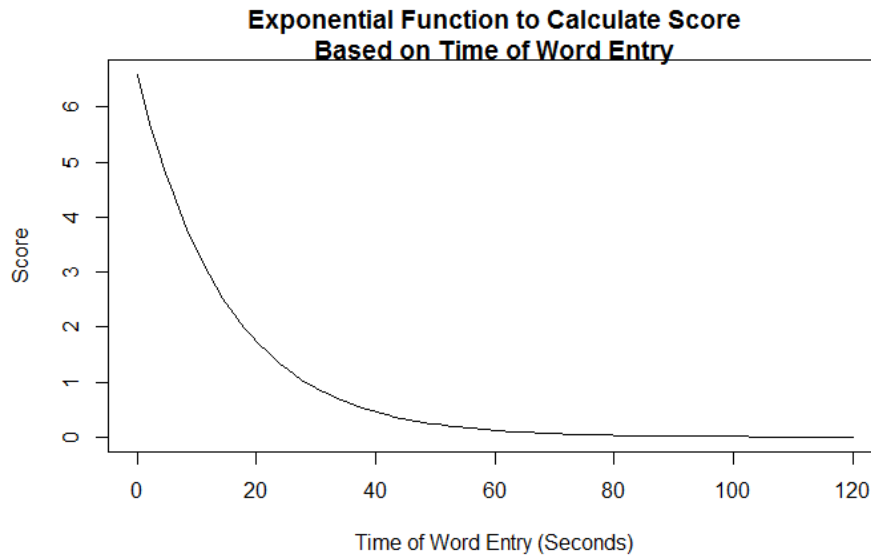


Figure 10: Time of Entry Score Function - Words entered in the first 20 seconds of the entry window receive the most weight.

Approach 4 uses the same exponential algorithm adjusted to calculate a score for each participant within each group, giving similar weight to the word time of entry and word score. While the word score provides the accuracy of a participant’s episodic memory, time of word entry information provides a measure of a participant’s episodic memory recall speed. Combined they can provide a more complete measurement of episodic memory by accounting for accuracy and speed.

Measure 5 – Dot Matrix T5 (Visual spatial working memory [15]):

Data Cleaning: Because all participants see the same number of dot matrix equations and grids, there is no pre-processing necessary.

Analysis Technique: For each participant, a confusion matrix is generated such as the one shown below in Table 7 for the Dot Matrix equation results. From this accuracy

is calculated for each participant per version. The Dot Grid accuracy is calculated by taking the number of correctly identified dots and dividing them by the number of total dots shown for a set. A second approach will be down that gives partial credit for incorrect dots that are close to the dot. A paired Wilcox test will be conducted to determine if there is a significant difference between the treatment and control groups

Table 7: Example Confusion Matrix for Dot Matrix Equation

	Is True	Is False
Response True	19	3
Response False	1	22

Treatment – Hand/Finger Movement T10 (Reaction time):

Data Cleaning: The software used to run the treatment calculates an activity rate per frame. The activity rate takes into account the number of clicks the participants has made with the mouse as well as the number of key presses. Thus the activity rate gives an empirical observation of a participant’s activity. This rate can be averaged per participant for use in analysis.

Analysis Technique: The average activity rate per participant is collected. A mid-point is selected that places the participants into two groups, high and low activity. Both groups have 8 participants. Using data from the Reaction Time Test task, the average reaction time of each group is looked at. Then a two sample, one-tailed Wilcox test is used on both high and low activity groups to determine if the mean reaction time in the treatment group is significantly more than mean reaction time of the control group. A Cliff’s Delta is also calculated on the high and low group data in order to determine the

effect size. The results provide evidence as to the level of significance and size of effect for both high and low groups.

Summary

This research includes several different research questions as outlined in Table 2. To address these research questions an alternate hypothesis is developed for each one as outlined in Table 3. To test each alternate hypothesis several dependent measures were developed. The dependent measures are each developed through cognitive measurement activities. These measurement activities make up the experiment. To determine if hand/finger movement prior an activity does have an effect on cognitive functions, a treatment and control version of the experiment is used to detect differences and evaluate all hypotheses. All data has been collected and analyzed for the stated purposes. The next chapter discusses those results.

IV. Analysis and Results

Chapter Overview

The experiment conducted for this thesis consisted of five individual activities measuring different cognitive functions. A treatment and control activity was used to provide a measurable effect from hand/finger movement on those cognitive functions. There were 16 participants and each participant experienced both the treatment and control conditions. This chapter provides the results for that experiment which is described in the previous chapter. Treatment and control groups means and standard deviations (SD) were calculated and reported for each. A two-sample one-sided Wilcoxon test was conducted to determine if the treatment group saw a benefit from the treatment activity. The p -values reported offer a determination as to whether the difference between the means is significant. A Cliff's Delta value is calculated for each cognitive measurement activity but will only be applicable where the difference between the means is significant. These main results are shown in Table 8 but will also be described in each section.

Table 8: Result of Means - Means (and standard deviation) of cognitive variables for treatment of control groups (last row is for day 2 and day 3 groups). * indicates significant at alpha = 0.05. RT is reaction time, Cat is categorization, ms is milliseconds, and Acc is accuracy.

Activity	Cog Function	Treatment Mean(SD)	Control Mean(SD)	Units	p-Value	Cliff's Delta
RT Test	RT	337(60)	345(60)	ms	<0.001*	0.1150
Oddball	Cat	4(4.9)	1.38(1.7)	Errors	0.8219	0.0938
Oddball	RT	419(110)	428(80)	ms	<0.001*	0.1096
Go/NoGo	Cat	6(5.4)	6.6(4.3)	Errors	0.2553	0.2553
Go/NoGo	Impulse	0.75(2.0)	0.94(1.5)	Errors	0.3345	0.1367
Go/NoGo	RT	480(95)	485(107)	ms	0.1460	0.0001
Word(1)	Episodic	7.625(3.324)	6.688(2.39)	Words	0.0994	0.1797
Word(2)	Episodic	6.521(3.33)	5.208(2.47)	Words	0.0736	0.3086
Word(3)	Episodic	12.13(4.47)	10.07(4.04)	Score	0.1156	0.2188
Word(4)	Episodic	38.2(13.87)	32.08(11.44)	Score	0.1261	0.1484
Dot Grid 1	Working	91.2(18)	92.7%(18)	% Acc	0.8813	0.0527
Dot Grid 2	Working	94.6(11)	95.6%(11)	% Acc	0.8900	0.0535
Dot Eq	Working	97.9(2)	97.9%(3)	% Acc	0.5178	0.0020
Dot Both	Working	94.5(7)	95.3(7)	% Acc	0.8299	0.1556
		Day 2 Mean(SD)	Day 3 Mean(SD)			
RT Test	RT	0.339(0.06)	0.343(0.06)	Sec	0.9991	0.0436

Experiment Timing and Latency Considerations

In the reaction time results discussed in the later sections, the differences between the means of the control and treatment group reaction times are 11 ms or less. USB keyboards as used in this test can have a 5-25 ms latency. The screen used has a refresh rate of 5.85 ms. There are steps that can be taken to achieve 1 ms accuracy with the PsychoPy software, such as ensuring synchronization of page-flips and using additional hardware to retrieve additional information on responses for purposes of handling system variability. Though many common PsychoPy practices for precision were followed, the

ones mentioned above were not. It is therefore a possibility that any findings found could be a result of the system latency. However due to the number of observations, particularly in the Reaction Time Test, and the use of the same machine for all participants, the effects seen should still be considered. Being on the same machine, even if the latency varies it should stay within the same mean for both the treatment and control condition. Having over 2300 observations per condition does allow any variations to cancel each other out.

Reaction time Results & Discussion – Reaction Time Test (T1)

The Reaction Time Test activity required participants to hit the ‘space’ bar when a black ‘X’ appeared in the center of the screen. Once the ‘space’ bar is pressed, the ‘X’ will disappear and be followed by another ‘X’ after a random interval of time. The main purpose of the Reaction Time Test was to determine if reaction time improved with prior hand/finger movement (T1). The independent variable used in this task was whether the participant performed the hand/finger movement prior the task. Per participant, the dependent variable consisting of the amount of time it took to respond to each stimuli was recorded for each the control and treatment experiments.

With the test version of treatment or control as the main factor, a two sample one-tailed Wilcox test is used to determine whether the mean of the reaction times in the treatment condition is significantly less than the mean of the reaction times in the control condition. A Wilcox test was used because the distributions of the reaction times were not normal, but skewed right. This right skewness is due to the nature of the activity.

Participants are trying to press a key as quickly as they can, but because of physical limits, people cannot react much faster than 0.2 seconds. However, there is no limit on the other side, meaning a person can be as slow to react as they want. In general, participants try to react quickly and the resulting distribution of reaction times is skewed right. Most values fall within 250 ms to 450 ms. The results of this analysis can be viewed in Table 8. When participants performed the treatment activity, they performed significantly faster ($p < 0.001$) in the Reaction Time Test activity than they did after performing the control activity. The mean reaction time for the treatment group is 337 ms, vs the control group's mean reaction time of 345 ms. A Cliff's Delta test was performed to provide an estimate of the size of effect the treatment activity had. The Cliff's Delta value of 0.115, which is a small effect. The difference between the means of the two groups is 8 ms. This is to say that on average the treatment group performed 8ms seconds faster than the control group. This improvement in reaction time may be of importance to specific operator fields, especially at the small time cost of performing a hand/finger movement for 80 seconds.

The combined evidence discussed above supports the rejection of measure 1's null hypothesis that the mean reaction time of a participant performing the treatment condition is equal to the mean reaction time of the same participant performing the control condition. Thus at the 95% confidence level, the results from these Wilcoxon tests further support the rejection of measure 1's Null hypothesis in favor of the alternate hypothesis: that the treatment reduced reaction time (T1).

Testing for Presence of a Learning Effect – Reaction Time Test (T9)

A two sample one-tailed Wilcoxon test was used to determine if the groups mean reaction time on day 3 is statistically less from the same group's mean reaction time on day 2, which would mean a learning effect is present. These results can be viewed in Table 8. Participants performed significantly faster ($p = 0.0153$) in the Reaction Time Test activity on day 2 than they did on day 3. The mean reaction time for the day 2 group is 339 ms, vs the day 3 group's mean reaction time of 343 ms. A Cliff's Delta test was performed to provide an estimate of the size of effect day 2 had. The Cliff's Delta value is 0.0436 which is an extremely small effect. This means an effect is present, but it is not a learning effect because the reaction time for the day 2 group is faster than the day 3 group. We would need to see that day 3 reaction time was faster than day 2 reaction time in order to have adequate evidence to suggest it a learning effect. This degradation in reaction time is possibly due to the participants become more relaxed and familiar with the experiment process by day 3.

The evidence discussed above supports the rejection of the null hypothesis that a learning effect is present from day 2 to day 3. Thus at the 95% confidence level, the results from these Wilcoxon tests further support the rejection of the Null hypothesis in favor of the alternate hypothesis: that a learning effect is not present from day 2 to day 3 for reaction time (T9).

Diminishing Effect Size – Reaction Time Test (T8)

The above results show that the treatment significantly reduces reaction time of participants. The next phase of analysis determines if the effect of the treatment diminishes over time. A sliding window method is used to perform a Cliff's Delta calculation on different sets of data, sliding across the duration of the activity. Different window sizes of 20, 40, 60, and 100 observations from each condition (treatment and control) were evaluated. Each window overlapped the preceding window by its window size minus 1. For example, a 100 observation window overlaps 99 observations in the previous window's data. Amongst the varying sizes used, the results across all sizes suggested the same conclusions. Plots were made for visual inspection. When considering all data, without controlling for any factor, the sliding window method gives us the plot as shown in Figure 11, which shows that the effect size seems to diminish over the course of the activity. A Wilcoxon test of significance between the first half and second half of the timeline confirms what is visually seen, that the effect of the treatment significantly ($p < 0.001$) diminishes over the duration of the activity.

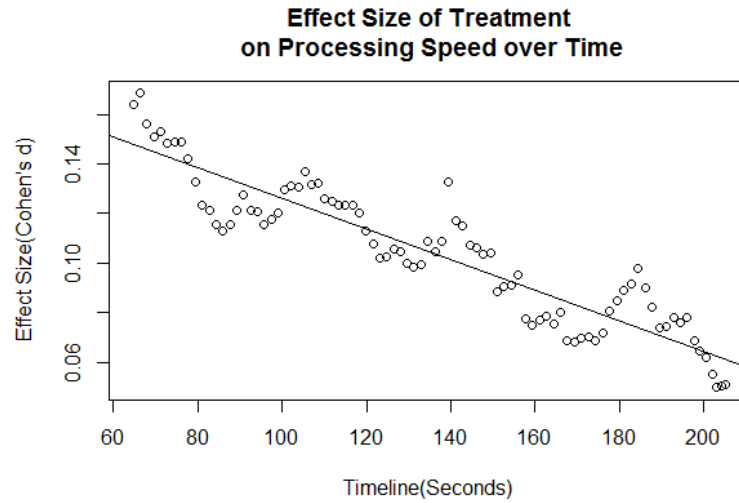


Figure 11: Effect Size of Treatment - Each dot represents the Cliff's Delta value for a 60 observation window of participant data within the duration of the Reaction Time Test activity.

The same sliding window method was used for an inspection of average reaction times over time. For Figure 12, a 40 observation window was used to average reaction times and plot them.



Figure 12: Reaction Time Averages over Time - Each dot represents the average reaction time for 40 observation window of participant data within the duration of the Reaction Time Test activity. Condition 1 is Control, condition 2 is Treatment.

Visually we see that control the group (condition 1 in Figure 12) started with a mean reaction time at approximately 340 ms and ended with a mean reaction time at approximately 353 ms, a difference of 13 ms. Because the control has no treatment effect which could may wear off, the reduction in reaction time over the course of the Reaction Time Test is believed to be due to fatigue. The treatment group (condition 2 in Figure 12) started with a mean reaction time of approximately 326 ms and ended with a mean reaction time at approximately 351 ms, a difference of 25 ms. The steeper slope is believed to be the effects due to fatigue with the additional effect of the treatment, which wears off almost completely by the end of the Reaction Time Test as the differences in mean reaction times at the end is only approximately 2 ms, versus the 14 ms difference at

the start. The results shown in Figure 12 support those in Figure 11 that the treatment does produce an effect, which diminishes over time.

The sliding window method used to determine effect size over time was also applied to different factors to gain more insight on demographic information. Four demographic factors were used with each factor being split into two categories with a similar number of participants: Age at which the participant started playing video games (Figure 13: 4-12 years “preteens” vs 13-38 years “teens or later”), Age of participant (Figure 14: 21-25 years vs 26-38 years), How many years the participant has played video games (Figure 15: 0-15 years vs 16+ years), and number of days a week the participant exercises (Figure 16: 1-3 days vs 4-6 days). While the plots show results for a particular window size, the intercept and slopes reported next were from all window sizes considered averaged for each finding and thus the numbers reported for the following discussion may not directly match the plots, however they provide the same end result. The plots use a polynomial (x^2) model to fit the data points and include grey confidence bound at 95%, meaning only 5% of data points are predicted to fall outside the grey confidence bounds. Referring to Figure 13, even though participants who started playing at 4-12 years of age saw an initially larger effect (Cliff’s Delta 0.245 vs 0.172), the effect diminished at a rate faster rate which resulting in the treatment having a more significant ($p = 0.001$) effect on participants who started video games in their teens and older.

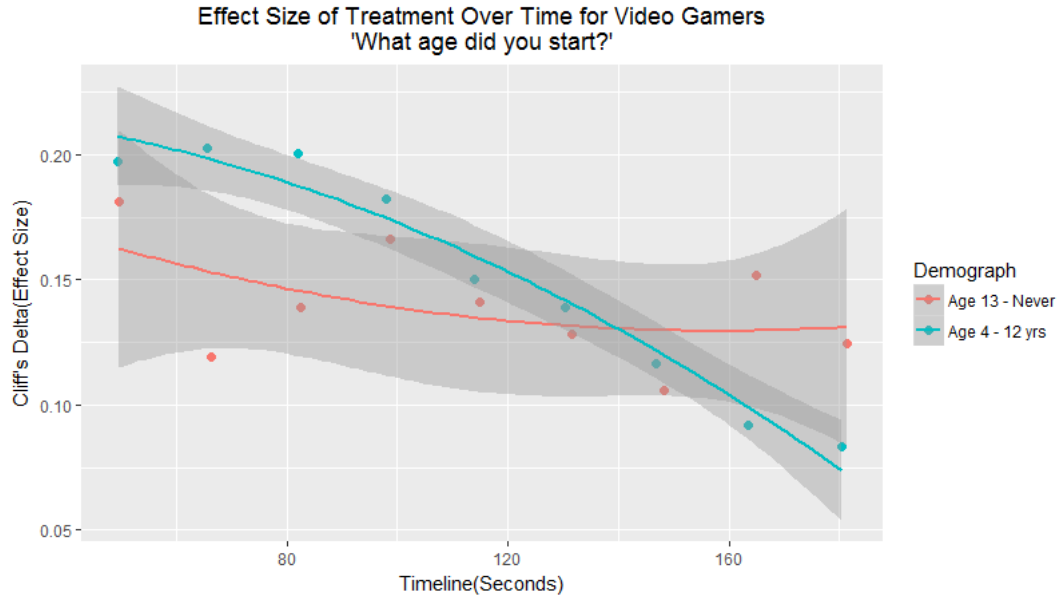


Figure 13: Effect Size of Treatment for Age When Started Video Games - Preteens consists of ages 4-12 years. Teens or later consists of ages 13-38 years. Both groups start with similar effect sizes but the effect diminishes quicker for preteens. Each dot represents the Cliff's Delta value for a 60 observation per participant, per condition window of within the duration of the Reaction Time Test activity.

Participants who have played video games for 16 years or more were significantly ($p < 0.001$) less affected by the treatment than those who have not played as long.

Though both groups started with a similar effect size, as seen in Figure 14, the diminishing rate of the effect is greater for participants who have played video games for 16 years or more. For this group, at approximately 125 seconds into the Reaction Time test, the effect switch from increasing performance to decreasing performance.

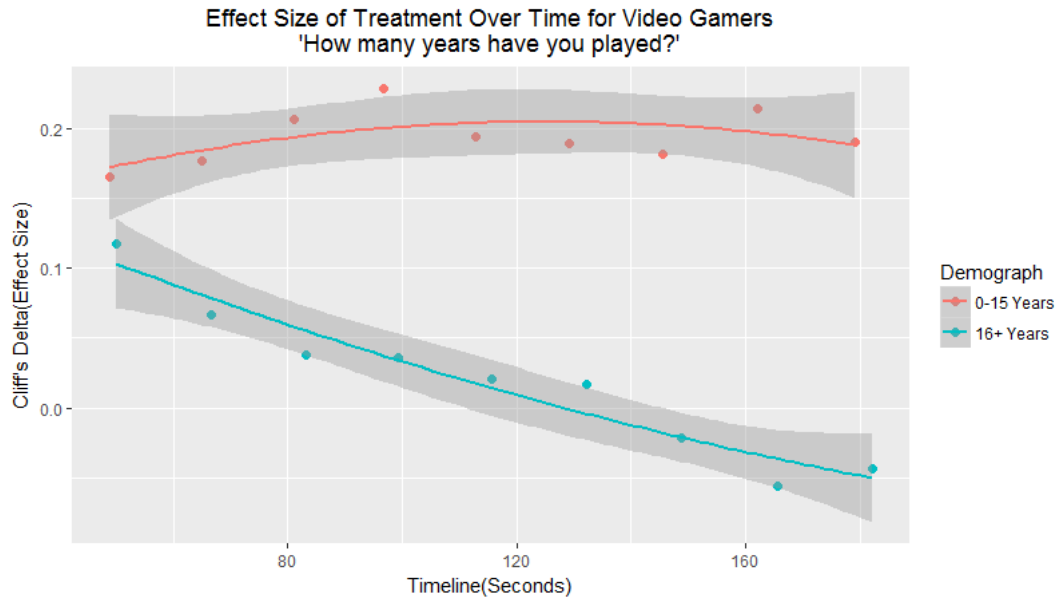


Figure 14: Effect Size of Treatment for Years Video Games Played - The effect of treatment diminishes at a similar rate for both, but participants with less video gaming years saw a larger effect.

Referring to Figure 15, participants who are 26 to 38 years of age were significantly ($p < 0.001$) more affected by the treatment than those who were 21 to 25 years of age. The two groups started around the same effect size, but after 120 seconds into the Reaction Time test, the effect size decreased for participants 21 to 25 years of age and increased for the participants 26 and older.

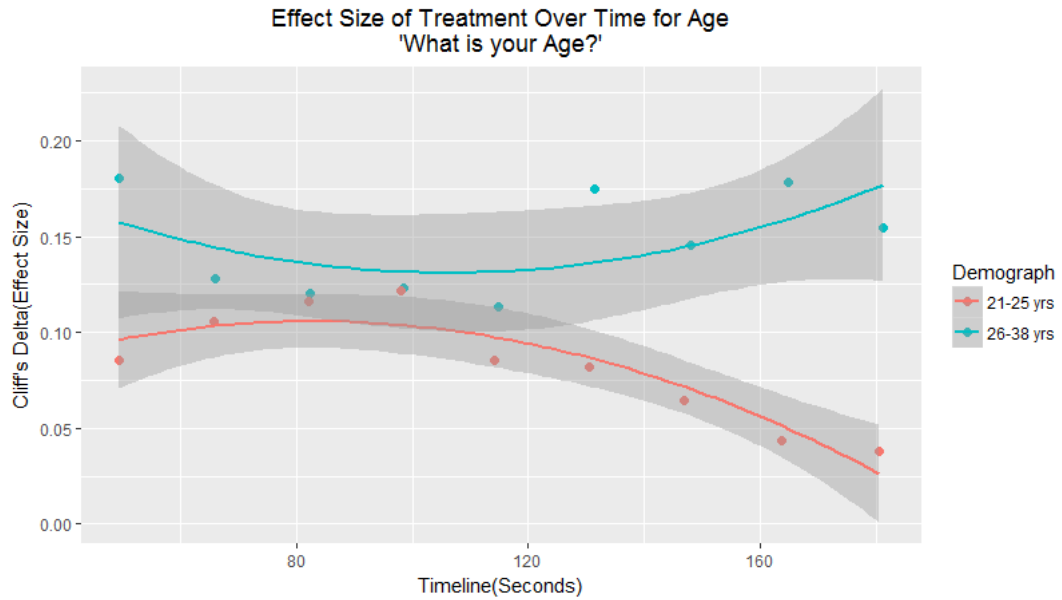


Figure 15: Effect Size of Treatment for Age - Both groups start with similar effect sizes but the effect diminishes quicker for 21-25 year group.

The demographics discussed thus far suggest that the effect of the treatment will be greater for participants who are older, less experienced, and who started playing video games later in life. For the final demographic and referring to Figure 16, participants who exercise four to six times per week are significantly ($p < 0.001$) more affected by the treatment than participants who exercise one to three times per week. Both groups have a similar effect size at the beginning and end of the Reaction Time test.

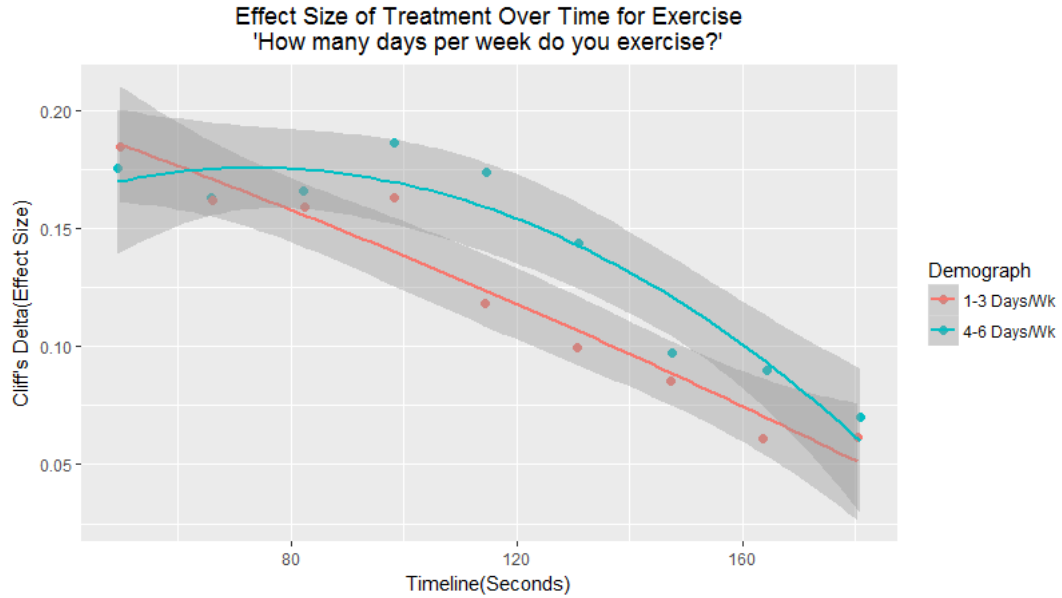


Figure 16: Effect Size of Treatment for Exercise Frequency - The effect is higher for participants who exercise more often but it diminishes faster than those who exercise less.

The combined evidence discussed above supports the measure one's null hypothesis that effects of the treatment activity of reaction time will diminish over time at the 95% confidence level (T8). Fatigue is present and does account for some diminishing of reaction time over the course of the Reaction Time Test, however the evidence above strongly supports the presence of an effect from the treatment that is in fact diminishing over time.

Treatment Activity Rate and Size of Effect – Reaction Time Test (T10)

During the treatment activity, participants moved and clicked the mouse with their right hand and repetitively pressed two keys with their left hand, simulating movement similar to personnel computer gamers. This information was used to calculate an average

activity rate per participant. By experimental design, participants were prompted to maintain activity rates between 0.6 and 0.8. Across all participants, the average activity rate was 0.66 with a standard deviation of 0.3. This is a large spread, and a 0.625 midpoint was selected to divide the participants into two equal size groups, a high and low activity group.

Initial inspection of the data showed that the high activity rate group has a 320 ms mean reaction time in the Reaction Time Test activity, while the low activity rate group has a 354 ms mean reaction time. On both the high and low activity groups, a two-sample, one-tailed Wilcoxon test was used to determine if the treatment group had a significantly faster reaction time than the control group. The results showed a larger p -value for the higher activity group than the lower activity group ($p = 1.055e-12$ vs $p = 5.133e-7$ respectively). The size of the treatment effect was larger for the higher activity group than the lower activity group (Cliff's Delta = 0.168 vs Cliff's Delta = 0.088 respectively).

A correlation coefficient was calculated between the activity rate and the size of the treatment effect. Although the correlation coefficient was not significant ($p = 0.4594$) with a correlation of 0.1992, it suggests that as the activity rate increases, the treatment effect also increased. However, one participant, we will call participant X, did not fit the trend of the rest of the data. This participant had the lowest activity rate at 0.27 a high Cliff's Delta value of 0.3831. The mean effect among all participants was 0.1333 with a standard deviation of 0.1915. Thus participant X's effect size falls outside one standard

deviation by 0.3. Participant X had the highest handicap multiplier. While 5 total participants had handicap multipliers, Participant X's multiplier was equivalent to the sum of the other four participant multipliers. This means with little effort, participant X was able to visually achieve the desired marker position in the treatment activity described in Chapter 3. Participant X is the oldest non-video game player by 11 years, as well as the oldest participant by 6 years. This outlier with unique demographics may be an indicator as to how a large group with similar attributes would perform as well as what the effect of the treatment may be on such participants. However, without other similar participants it is impossible to determine if this effect is due to the participants demographics or if it was a unique case. Thus for purposes of a determining a representative correlation coefficient, participant X was removed from this calculation, producing a coefficient for participants age 21-33, which consist of two non-video gamers. The new correlation coefficient with 15 of the 16 participants considered is now a moderately strong positive coefficient at 0.5397, and is significant at $p = 0.0379$. Participant X was left in for all other analysis as while there was a slight decrease in p -values in Wilcox tests, the removal would not change any conclusions.

The evidence above suggest we reject the null hypothesis that greater hand/finger movement during the treatment activity does not increase the size of the effect of treatment. The evidence suggest the alternative hypothesis: that the greater the hand/finger movement during the treatment activity, the greater the effect size of the treatment.

Categorization Results & Discussion – Oddball and Go/No Go (T2)

Both Oddball and Go/No Go tasks were used to determine if categorization error is reduced with prior hand/finger movement. Participants pressed the right or left control key when the stimulus was shown as directed. The independent variable used in this task was whether the participant performed the hand/finger movement prior the task. The number of identified stimuli (dependent variable) was recorded for each the control and treatment experiment versions. The number of stimuli incorrectly categorized was tallied and error was calculated. The following is the result of that analysis.

With the test version of treatment or control as the main factor, a two sample one-tailed Wilcox test was used to determine if the group's categorization number of errors after performing the treatment activity is statistically reduced from the same group's categorization number of errors after performing the control activity. These results can be viewed in Table 8. When participants performed the treatment activity, they did not perform significantly better ($p = 0.8219$ for Oddball, $p = 0.2553$ for Go/NoGo) in the Oddball and Go/NoGo measurement activities than they did after performing the control activity.

Again with the test version of treatment or control as the main factor, another two sample one-tailed Wilcox test was used to determine if the group's reaction time after performing the treatment activity is statistically better from the same group's reaction time after performing the control activity. When participants performed the treatment activity, they performed significantly better ($p < 0.001$) in Oddball, but not significantly

better ($p = 0.1460$) in Go/NoGo than they did after performing the control activity. The effect size for Oddball is small (Cliff's Delta = 0.1096)

The combined evidence discussed above does support the rejection of null hypothesis that the mean reaction time of a participant performing the treatment experiment is equal to the mean reaction time of the same participant performing the control experiment. Thus at the 95% confidence level, the results from these Wilcox tests further support the rejection of the Null hypothesis in favor of the alternate hypothesis: that the treatment reduced reaction time (T1).

Inhibitory/Impulse Control Results & Discussion - Go/NoGo (T3)

The Go/NoGo activity was used to determine if inhibitory/impulse control number of errors is reduced with prior hand/finger movement. The independent variable used in this task was whether the participant performed the hand/finger movement prior the task. The number of stimuli there were ignored, that were supposed to be ignored (dependent variable) was recorded for each the control and treatment experiment versions. The number of errors was tallied. The following is the results of that analysis.

With the test version of treatment or control as the main factor, a two sample one-tailed Wilcox test was run on the treatment and control groups. The purpose of the Wilcox tests was to determine if the group's inhibitory/impulse control error count after performing the treatment activity is statistically better than the same group's categorization error count after performing the control activity. These results can be

viewed in Table 8. When participants performed the treatment activity, they did not perform significantly better ($p = 0.334$) in the Go/NoGo measurement activity than they did after performing the control activity.

The evidence discussed above does not support the rejection of measure 3's null hypothesis that the "inhibitory/impulse control error of a participant performing the treatment experiment is equal to the mean inhibitory/impulse control error of the same participant performing the control experiment." Thus we cannot claim that the treatment reduced inhibitory/impulse control error in the Go/NoGo activity (T3).

Episodic Memory Results & Discussion – Word Recall (T4)

The word recall task took place in two phases with the Oddball task in between the two phases, allowing a 210 seconds distraction before the second phase. During the first phase the participant views 30 unique 6-7 letter words. The words are displayed individually for 1 second followed by a 0.25 second break between words. In the second phase participants are asked to type as many words as they can recall within 120 seconds. The main purpose of the Word Recall task was to determine if episodic memory improved with prior hand/finger movement. The independent variable used in this task was whether the participant performed the hand/finger movement prior the task. The dependent variables are the number of words recalled and the time at which those words were submitted. The number of words recalled per participant per group was tallied as well as individually scored based on spelling errors and relation to actual word. In addition, an exponential algorithm was used to give more weight to words entered earlier

in the second phase as discussed in Chapter 3 (pseudo code provided in Appendix E). Analysis was conducted with four different approaches altering the use of scoring and time of entry information. The following are the results of that analysis.

Table 9: Scoring Approaches for Word Recall

Approach	Formula Description
1	Only words that were a direct match were counted
2	Words scoring method applied: Exact Match = 3 1 or 2 errors = 2 2 or 3 errors = 1 Wrong word but same category as goal word = 0 No relation to goal word = -1
3	Words given a score based on time of entry, heavily rewarding words entered closer to the start of the word entry window. Described in Chapter 3.
4	Word score from approach 2 and time score from approach 3 are added together. 'w' in Equation 2 is used to give word and time scores equal weight.

With the test version of treatment or control as the main factor, a two sample one-tailed Wilcoxon test was run on the treatment and control groups for each of the four approaches. The purpose of the Wilcoxon tests was to determine if the group's episodic memory after performing the treatment activity is statistically better than the same group's episodic memory after performing the control activity. The four approaches allow for the individual inspection of the timing and correctness for each groups episodic memory as well as the combined effects. These results can be viewed in Table 9 with the approaches labelled as 1-4. For approach 1, considering only the unscored word information, participants did not perform significantly better ($p = 0.0994$) in the Word

Recall measurement activity than they did after performing the control activity. In approach 2, where word scoring information was included for purposes of partial credit, the argument for significance becomes more substantial as the p value drops to 0.0736. For approach 3, which considered only time of entry information, the p value is 0.01156. Adding together the word score and time of entry information, with similar weights for both, produces a p value of 0.1261, indicating participants did not perform significantly better in the Word Recall activity in the treatment condition than they did after performing the control activity.

The evidence discussed above does not support the rejection of measure 4's null hypothesis that the "mean episodic memory accuracy of a participant performing the treatment experiment is equal to the mean episodic memory accuracy of the same participant performing the control experiment. (T2)" The evidence discussed also does not support the rejection of measure 4's null hypothesis that the mean episodic memory recall speed of a participant performing the treatment experiment is equal to the mean episodic memory accuracy of the same participant performing the control experiment (T6). However, the low p values are low enough to warrant further studies in this area.

Visual spatial working memory Results & Discussion – Dot Matrix (T5)

The Dot Matrix measure was used to determine if visual spatial working memory improved with prior hand/finger movement. The independent variable used in this task was the whether the participant performed the hand/finger movement prior the task. The number of correctly identified matrixes (dependent variable) and the number of correctly

identified dots (dependent variable) were recorded for each the control and treatment tasks. The number of correctly and incorrectly identified matrices and dots were tallied. The following are the results of that analysis.

With the test version of treatment or control as the main factor, a two sample one-tailed Wilcoxon test was run on the treatment and control groups for the Dot Matrix Grid and Equation results separately, and then together. The purpose of the Wilcoxon tests was to determine if the group's classification accuracy after performing the treatment activity is statistically more accurate than same group's visual spatial working memory accuracy after performing the control activity for the three categories (Dot Grid, Dot Equation, and Both Together). These results can be viewed in Table 8. The results show that when participants performed the treatment activity, they did not perform significantly better (Grid: $p = 0.8813$, Equation: $p = 0.5178$, Both: $p = 0.8299$) in the Dot Matrix measurement activity than they did after performing the control activity for all three categories. Instead they performed worse, though not statistically significant (Grid: $p = 0.1199$, Equation: $p = 0.5178$, Both: $p = 0.1767$). Adding partial credit for dots close to the missed dots further reduced the difference between the participant performance in the treatment and their performance in the control condition. The results show that participants did not perform significantly better in the treatment condition ($p = 0.8900$). However they did not perform significantly worse ($p = 0.1431$).

The evidence discussed above does not support the rejection of measure 5's null hypothesis that the "visual spatial working memory accuracy of a participant performing

the treatment experiment is equal to the mean visual spatial working memory accuracy of the same participant performing the control experiment.” Thus we cannot claim that the treatment increased or decreased visual spatial working memory accuracy in the Dot Matrix activity (T5). One fact to consider is that 169 of 480 observations were 100% accurate for both groups, meaning those observations were not useful. This issue is discussed further in Chapter 5

Summary

The results showed that when participants performed the treatment activity, their mean reaction time was significantly faster ($p < 0.001$ Reaction Time Test, $p < 0.001$ Oddball) than the same participants’ mean reaction time after performing the control activity in Reaction Time Test, Oddball, and Go/NoGo. The effect size is small as measured using Cliff’s Delta for all three cases. A small but significant anti-learning effect is present between day 2 and day 3 for the Reaction Time Test. This is to say that the opposite of a learning effect is occurring, thus instead of seeing an increase in performance on day 3, a significant decrease was found. Participants who performed the treatment activity achieved a non-significantly ($p = 0.0736$) higher score in the Word Recall activity than the control group in scoring approach 2. p -values for Go/NoGo categorization and Word Recall episodic memory scoring approaches 1-3 are low enough to elicit further research. All other cognitive measurement activities were unable to show a significant effect due to the treatment activity. For some measurement activities, the

means of the control group were better, though not significant. The next chapter covers possible future work in this area, some lessons learned, and conclusions.

V. Conclusions and Recommendations

Chapter Overview

The experiment for this thesis evaluated whether hand and finger movement prior to a task affects a person's cognitive performance. To evaluate that hypothesis, five cognitive measurement activities were designed and conducted, each with their own hypotheses. This chapter outlines the findings of each measure in regards to their individual hypotheses, and the final evaluation of the main hypothesis. Following said discussion, recommendations will be addressed. These recommendations include improvement and suggested improvements for the current test design, as well as different related avenues future research could take.

Conclusions of Research

On day 2 or 3 of the experiment, participants either performed a treatment or control activity four times of which was interwoven with the five cognitive measurement activities. The data collected from these cognitive measurement activities was analyzed. The mean of each group (treatment and control) was calculated reported in Table 8 located in Chapter IV. A two-sample one-sided Wilcox test p -value to determine significance between the means was calculated as well as Cliff's Delta in order to determine the size of the effect if present. For the Reaction Time Test and Oddball activities, the results show that participants who performed the treatment activity had a mean reaction time that was significantly faster ($p < 0.001$, $p < 0.001$ respectively) than

the same participants' mean reaction time after performing the control activity. Cliff's Delta value came to 0.115 and 0.1096 respectively, indicating relatively small effects. There was a concern that a learning effect was present in the Reaction Time Test activity and thus similar tests were run but with the participants grouped by day 2 or day 3. A learning effect would mean reaction time is slower on day 2, however a significant anti-learning effect was found, with a p -value of 0.0153 and a Cliff's Delta of 0.045. This effect was significant but much less than that of the treatment, also the effect was less than half that of the treatment. It is believed the tediousness of the experiment may have contributed to this degradation in reaction time. Using a scoring method for the Word Recall activity which considered both time and correctness, a significant ($p = 0.0436$) increase in the treatment group was found.

All other cognitive measurement activities were unable to show a significant effect at $\alpha = 0.05$ due to the treatment activity. The word recall scoring approaches 1 through 4 and oddball cognitive measurement activities had p values between 0.07 and 0.13. While this indicates the difference in the means are not significant, it is low enough to elicit further research taking advantage of lessons learned which will be discussed further in this chapter.

This experiment has shown that hand/finger movement does have a significant beneficial effect on the cognitive function of reaction time (T1) and episodic memory when considering timing information. This supports the claim that the main alternate hypothesis can be accepted: that there is a cognitive effect present from hand and finger

movement.

Significance of Research

This thesis focuses on the small-muscles within the hand and fingers and shows there is some correlation between the movement of those muscles and cognitive improvement. The benefit to the scientific community is the unique methods of exploration used in a highly uninvestigated area. This pilot study provides a foundational building block for future research in this focus area and an expansion of research already done. The study begins to answer the question of what can be done to help our military operators when their cognitive functions are in a reduced state.

Recommendations for Action & Future Research

Being a pilot study, there are many lessons learned for future experiments. Though some of these lessons came about by pure realization of the issue as the experiment progressed, many of these risks or testing imperfections were known before testing, being accepted due to time or resource constraints. Below the main lessons learned are discussed in detail with the general theme focusing on improving the current test design.

First, as designed most cognitive measurement activities used do not provide the sensitivity within the data that they could due to large numbers of non-unique observations between treatment and control. Compare and contrast the Reaction Time Test activity with the Dot Matrix activity. The Reaction Time Test measure provides a

wide distribution of response times for each reaction time. For example, the treatment group's mean was 337 ms, but participants could go as low as near 100 ms (physical limit) to above 1000 ms or more. Within that distribution are 16,000 unique response times a participant's reaction time could be at. Reaction time tests also have the benefit that people vary significantly in their reaction times. Thus reaction time activities produce a list of results that is well distributed with nearly 99% of the 5000 observations being unique. This provides sensitivity for the detection of an effect. The Dot Matrix Equation accuracy mean for the treatment group was 96.6%, with the lowest accuracy at 93.3%. Thus 50% of the error rate is at 96.6% accuracy. The closeness of the results came from the difficulty, or lack thereof, in the measure, creating a ceiling effect. Of the 16 participants in the treatment group for the dot matrix equations, 7 got perfect scores. While that is good for the participant, a perfect score among 7 participants leaves little room for analysis and determining effect. A suggested redesign would include making the Dot Matrix measure more difficult so that the mean accuracy fell around 0.50. Making it more difficult would increase the number of unique observations. Increasing the equation matrix from 3x3 to 4x4, and increasing the dot location grid from 5x5 to 6x6 would be a possible design change. This should inherently increase the variation in results and provide much more sensitive data.

The second area for improvement is to further develop the small-muscle movement (warm-up) activity. The current design has several objectives:

1. Minimize activation of the brain's cognitive functions

2. Maintain similar hand and finger movement rates among participants
3. Reduce making the participant feel stressed or excited

If the warm-up activity engages the participant's cognitive functions in a similar fashion as the measures do, then the nature of the warm-up has changed from un-related to related with regard to the activities used to measure the cognitive functions. The warm-up activity in this thesis was designed to be challenging physically, but not mentally. Ideally, participants would be in a relaxed state and move their hands and fingers within some standard rate. This mindless activity is ideal for non-related physical hand/finger activities as the participant does not need to look at anything or think about anything except to continually move the fingers. However in an effort to measure and standardize the movement rate, the hand and mouse keyboard warm-up activity was created. This effort to measure and standardize the rate may have unknown confounding effects on cognitive functions. No longer are participants mindlessly moving their hands and fingers but are being persuaded to maintain the standard rate. This rate maintenance could cause undesired emotions to arise within certain participants such as stress, while giving the feeling of excitement to others. Feelings such as stress can effect cognitive functions [36]. The current warm-up activity achieved minimizing stress or feelings of excitement by the methods used to persuade participants to maintain their movement rate while limiting the induced stress and excitement. However there is much room for improvement in the design of the warm-up activity.

In addition to improving upon the current test design, there are several ways to

modify the experiment to explore other questions. The following is a discussion of those avenues and what they may show.

First, is the question of how long does the cognitive effect by small-muscle warm-up last. For gum chewing, this question was by Onyper et al. and was shown to remain significantly beneficial to some of the participant's cognitive functions for 15-20 minutes before wearing off [15]. But this thesis experiment design was intentionally not geared towards determining why the effect wears off, but instead maximizes the presence of the effect. Most measures were 4 minutes in length following the treatment, and the longest measure lasted approximately 12 minutes. If the effects of the treatment in this experiment lasted as long as they did in Onyper's, then the end of the effect could not be seen as all measures end too soon. Thus the true limits of the effect from hand and finger warm-up remain unexplored.

Second, this study focused on ensuring the warm-up task was not related to the tasks performed within each measure. However there is also a research gap in small-muscle movement and the effect on cognitive performance in related tasks. Medical studies are currently pioneering this research, but there are many fields other than just surgical that can benefit from this type of research [6]. More importantly, the medical studies tend to focus on related small-muscle warm-up and its effect on surgical performance, not cognitive performance. Because they focus on surgical performance, there is little ability to say which cognitive functions were improved by a related warm-up. Thus if a particular manager wants to help increase the reaction time of his or her

operators by developing a related warm-up activity, the surgical performance focused studies are not useful because they do not focus on particular cognitive functions. Any related small-muscle warm-up effects study that focuses on cognitive performance would be more applicable to the wide range of operator fields because the cognitive functions will be the same no matter what the field is.

Third, the cognitive functions examined in this thesis are just a few of many that may prove useful in measuring the cognitive effect of small-muscle movement. Some cognitive functions may be easier to measure than others. Others may not be affected by such small-muscle movement, such as verbal fluency in the gum chewing study done by Onyper et al. [10]. These types of questions remain unstudied in the realm of small-muscle warm-up. John Krantz provides a list of several cognitive functions and a basic test for measuring them on his Hanover College webpage [11].

Lastly, a possible confounding factor that may be present in this study is the unknown cognitive effect induced by each measure on the other subsequent measures. The small-muscle warm-up activity was completed before each individual measure; however, the delay between measures was approximately 2 minutes. From Onyper's study, it is known that the warm-up benefit seen from gum chewing lasted for about 15-20 minutes [15]. If a cognitive effect is caused by a measure in this thesis experiment, then it would be reasonable to assume that the effect would still be present for the next measure, even if the treatment had no cognitive effect. The main purpose of this thesis was to determine if a cognitive effect caused by the treatment was present at all. Thus the

order of the measure was purposely not varied as they were in Onyper et al. study, which showed the order in which the measures were given within the treatment or control condition had no significant effect on the results [15]. The purposeful decision to not vary the measures was made because if an effect does exist by varying the measures, it is not believed to be big enough to matter in regards to the purposes of this thesis experiment. However this effect could be explored by running a similar experiment and varying the measure sequence.

Summary

In summary, this thesis experiment shows that hand and finger movement prior to task does have a significant effect on cognitive performance. This was shown for the reaction time cognitive function, where a small effect was seen among the participants when performing the hand and finger treatment activity. The other cognitive measurement activities did not show a significant effect. This is either because there was none or cognitive measurement activities were not sensitive enough to provide meaningful differences. Further studies are needed to improve upon the techniques used in this experiment in order to further determine which other cognitive functions are effected by small-muscle movement and how.

Appendix A: Reaction time Results

A series of three two-sample paired Welch's t -tests were run on each participant's data from measure 1. Each participant's data was sectioned further by the delay time (0.5, 1.0, and 1.5 seconds) between stimuli before the t -tests were run, producing 48 rows of t -test results as shown in Table 10. The purpose of the three t -tests is to determine if a participant's mean reaction time after performing the control activity is statistically different from the same participant's mean reaction time after performing the treatment activity.

Table 10: Reaction time Results by Participant - Each row contains participant results for a particular delay. The results include three different *t*-tests

Participant ID	Delay	Two Sided T Test p-Value	(2-Sided) Fail to Reject HO	" C > T " T Test p-Value	(C > T) Fail to Reject HO	" T > C " T Test p-Value	(T > C) Fail to Reject HO	Day 2	Day 3
2396	0.5	0.8304	TRUE	0.4152	TRUE	0.5848	TRUE	T	C
2396	1	0.0675	TRUE	0.9662	TRUE	0.0338	FALSE	T	C
2396	1.5	0.1264	TRUE	0.0632	TRUE	0.9368	TRUE	T	C
2567	0.5	0.0876	TRUE	0.0438	FALSE	0.9562	TRUE	C	T
2567	1	0.5641	TRUE	0.2820	TRUE	0.7180	TRUE	C	T
2567	1.5	0.0812	TRUE	0.0406	FALSE	0.9594	TRUE	C	T
3577	0.5	0.0167	FALSE	0.9916	TRUE	0.0084	FALSE	C	T
3577	1	0.1075	TRUE	0.9462	TRUE	0.0538	TRUE	C	T
3577	1.5	0.1632	TRUE	0.9184	TRUE	0.0816	TRUE	C	T
4234	0.5	0.0046	FALSE	0.9977	TRUE	0.0023	FALSE	T	C
4234	1	0.6136	TRUE	0.3068	TRUE	0.6932	TRUE	T	C
4234	1.5	0.7979	TRUE	0.6010	TRUE	0.3990	TRUE	T	C
4277	0.5	0.2148	TRUE	0.8926	TRUE	0.1074	TRUE	T	C
4277	1	0.4048	TRUE	0.7976	TRUE	0.2024	TRUE	T	C
4277	1.5	0.2995	TRUE	0.1498	TRUE	0.8502	TRUE	T	C
4533	0.5	0.0046	FALSE	0.9977	TRUE	0.0023	FALSE	C	T
4533	1	0.5119	TRUE	0.2560	TRUE	0.7440	TRUE	C	T
4533	1.5	0.2347	TRUE	0.8826	TRUE	0.1174	TRUE	C	T
4600	0.5	0.0953	TRUE	0.9524	TRUE	0.0476	FALSE	C	T
4600	1	0.4467	TRUE	0.7767	TRUE	0.2233	TRUE	C	T
4600	1.5	0.0026	FALSE	0.0013	FALSE	0.9987	TRUE	C	T
4645	0.5	0.0126	FALSE	0.0063	FALSE	0.9937	TRUE	T	C
4645	1	0.0034	FALSE	0.0017	FALSE	0.9983	TRUE	T	C
4645	1.5	0.0092	FALSE	0.0046	FALSE	0.9954	TRUE	T	C
4953	0.5	0.1474	TRUE	0.0737	TRUE	0.9263	TRUE	C	T
4953	1	0.0077	FALSE	0.0038	FALSE	0.9962	TRUE	C	T
4953	1.5	0.0072	FALSE	0.0036	FALSE	0.9964	TRUE	C	T
6645	0.5	0.0353	FALSE	0.0176	FALSE	0.9824	TRUE	T	C
6645	1	0.0028	FALSE	0.0014	FALSE	0.9986	TRUE	T	C
6645	1.5	0.0000	FALSE	0.0000	FALSE	1.0000	TRUE	T	C
6952	0.5	0.0844	TRUE	0.0422	FALSE	0.9578	TRUE	C	T
6952	1	0.6289	TRUE	0.3144	TRUE	0.6856	TRUE	C	T
6952	1.5	0.2963	TRUE	0.1481	TRUE	0.8519	TRUE	C	T
7840	0.5	0.9221	TRUE	0.5390	TRUE	0.4610	TRUE	T	C
7840	1	0.7817	TRUE	0.3909	TRUE	0.6091	TRUE	T	C
7840	1.5	0.6494	TRUE	0.3247	TRUE	0.6753	TRUE	T	C
8334	0.5	0.0004	FALSE	0.0002	FALSE	0.9998	TRUE	T	C
8334	1	0.0031	FALSE	0.0016	FALSE	0.9984	TRUE	T	C
8334	1.5	0.0001	FALSE	0.0000	FALSE	1.0000	TRUE	T	C
8413	0.5	0.6872	TRUE	0.6564	TRUE	0.3436	TRUE	T	C
8413	1	0.0172	FALSE	0.0086	FALSE	0.9914	TRUE	T	C
8413	1.5	0.0050	FALSE	0.0025	FALSE	0.9975	TRUE	T	C
8968	0.5	0.0000	FALSE	0.0000	FALSE	1.0000	TRUE	C	T
8968	1	0.1574	TRUE	0.0787	TRUE	0.9213	TRUE	C	T
8968	1.5	0.0436	FALSE	0.0218	FALSE	0.9782	TRUE	C	T
9321	0.5	0.2847	TRUE	0.1424	TRUE	0.8576	TRUE	C	T
9321	1	0.3130	TRUE	0.1565	TRUE	0.8435	TRUE	C	T
9321	1.5	0.1416	TRUE	0.9292	TRUE	0.0708	TRUE	C	T
Total			19		19		5		

The second column in Table 10 represents the delay time between stimuli. Only reaction time data for that row's participant and delay time are considered in the *t*-tests performed. All null and alternate hypotheses for this *t*-test analysis are described in Table 11.

Table 11: Hypothesis's for Two-Sided *t*-tests - “ C > T “, and “ T > C “ with regards to reaction time of participants in measure 1. C = Control group. T = Treatment group.

Hypothesis Type	Hypothesis
Null	There is no significant difference between the means of the control and treatment groups
Two-Sided Alternate	There is a significant difference between the means of the control and treatment groups
“ C > T “ Alternate	There is significant evidence that the mean reaction time of the control group is greater than the treatments group's
“ T > C “ Alternate	There is significant evidence that the mean reaction time of the control group is less than the treatment group's

The next two columns show results from a “Two Sided” *t*-test. A p-value above 0.05 is labeled as “TRUE” in the second of the two columns considered, indicating we fail to reject the null hypothesis. A p-value below 0.05 is labeled “FALSE,” indicating the null hypothesis should be rejected in favor of the two-sided alternate hypothesis. The process described for reading these two columns can be directly applied the next set of two columns, which contains results from a one-sided *t*-test. This one-sided *t*-test specifically looks at whether the control groups mean reaction time is significantly greater than the treatment groups mean reaction time. The described method is further applied to the next two columns in Table 10. This one-sided *t*-test specifically looks at

whether the control groups mean reaction time is significantly less than the treatment groups mean reaction time. The final two columns indicate which day the treatment and control were done on, an important factor when determining the presence of a learning effect.

The results in Table 10 show that there is a significant difference between the mean reaction times for the control and treatment groups for 19 of the 48 participants. This two-sided test does not inform whether this difference was in favor of the control or treatment group (the “favored” group has a lower mean reaction time). 16 of the 19 participants where a significant difference was detected, the participants in the treatment group had a significantly faster reaction time, as indicated in the two columns for ‘C > T’ (C = Control group, T = Treatment group). The two ‘C > T’ columns also show three more results where the participants in the treatment group had a significantly faster reaction time, for a total of 19. The results in the ‘T < C’ show that five rows are in favor of the control group.

Appendix B: Learning Effect Individual Analysis

Because participants are their own control, it is possible that there is a learning effect occurring between the two days of testing. This possible effect was deterred by training all participants so that they became familiar with the tasks before testing began and by altering which day a participants did the treatment and control versions of the test. No measurement method was used to ensure participants had achieved a particular standard. Instead the training administer visually ensured each participant understood and could perform each activity proficiently. Given the precautionary steps however, a learning effect is still possible and is examined further for possible effect on reaction time. Table 12 is set up exactly the same as Table 10 in Appendix A; however the purpose of the three *t*-tests in this table is to determine if a participant's mean reaction time in on test day 2 is statistically different from the same participant's mean reaction time on test day 3. If a learning effect was present, it would be expected that the participant's mean reaction time in favors test day 3 ('D2 > D3'). Thus most results labeled "FALSE" should appear in the " 'D2 > D3' T Test *p*-Value" column. However 10 of 48 participants are found to have faster reaction times on day 3 and 14 participants have faster reaction times on test day 2 as indicated in columns for 'D3 > D2'. Unlike the results from Table 10 where 19 participants had faster reaction times in the treatment group and 5 participants in the control group had faster reaction times, these results indicate that there is not enough evidence to support the presence of a learning effect.

Table 12: Learning Effect Results by Participant - Each row contains participant results for a particular delay. The results include three different *t*-tests

participant	Delay	Two Sided T Test p_Value	(2-Sided) Fail to Reject HO	"D2 > D3" T Test p_Value	(D2 > D3) Fail to Reject HO	"D3 > D2" T Test p_Value	(D3 > D2) Fail to Reject HO	Day 2	Day 3
2396	0.5	0.8304	TRUE	0.5848	TRUE	0.4152	TRUE	T	C
2396	1	0.0675	TRUE	0.0338	FALSE	0.9662	TRUE	T	C
2396	1.5	0.1264	TRUE	0.9368	TRUE	0.0632	TRUE	T	C
2567	0.5	0.0876	TRUE	0.0438	FALSE	0.9562	TRUE	C	T
2567	1	0.5641	TRUE	0.2820	TRUE	0.7180	TRUE	C	T
2567	1.5	0.0812	TRUE	0.0406	FALSE	0.9594	TRUE	C	T
3577	0.5	0.0167	FALSE	0.9916	TRUE	0.0084	FALSE	C	T
3577	1	0.1075	TRUE	0.9462	TRUE	0.0538	TRUE	C	T
3577	1.5	0.1632	TRUE	0.9184	TRUE	0.0816	TRUE	C	T
4234	0.5	0.0046	FALSE	0.0023	FALSE	0.9977	TRUE	T	C
4234	1	0.6136	TRUE	0.6932	TRUE	0.3068	TRUE	T	C
4234	1.5	0.7979	TRUE	0.3990	TRUE	0.6010	TRUE	T	C
4277	0.5	0.2148	TRUE	0.1074	TRUE	0.8926	TRUE	T	C
4277	1	0.4048	TRUE	0.2024	TRUE	0.7976	TRUE	T	C
4277	1.5	0.2995	TRUE	0.8502	TRUE	0.1498	TRUE	T	C
4533	0.5	0.0046	FALSE	0.9977	TRUE	0.0023	FALSE	C	T
4533	1	0.5119	TRUE	0.2560	TRUE	0.7440	TRUE	C	T
4533	1.5	0.2347	TRUE	0.8826	TRUE	0.1174	TRUE	C	T
4600	0.5	0.0953	TRUE	0.9524	TRUE	0.0476	FALSE	C	T
4600	1	0.4467	TRUE	0.7767	TRUE	0.2233	TRUE	C	T
4600	1.5	0.0026	FALSE	0.0013	FALSE	0.9987	TRUE	C	T
4645	0.5	0.0126	FALSE	0.9937	TRUE	0.0063	FALSE	T	C
4645	1	0.0034	FALSE	0.9983	TRUE	0.0017	FALSE	T	C
4645	1.5	0.0092	FALSE	0.9954	TRUE	0.0046	FALSE	T	C
4953	0.5	0.1474	TRUE	0.0737	TRUE	0.9263	TRUE	C	T
4953	1	0.0077	FALSE	0.0038	FALSE	0.9962	TRUE	C	T
4953	1.5	0.0072	FALSE	0.0036	FALSE	0.9964	TRUE	C	T
6645	0.5	0.0353	FALSE	0.9824	TRUE	0.0176	FALSE	T	C
6645	1	0.0028	FALSE	0.9986	TRUE	0.0014	FALSE	T	C
6645	1.5	0.0000	FALSE	1.0000	TRUE	0.0000	FALSE	T	C
6952	0.5	0.0844	TRUE	0.0422	FALSE	0.9578	TRUE	C	T
6952	1	0.6289	TRUE	0.3144	TRUE	0.6856	TRUE	C	T
6952	1.5	0.2963	TRUE	0.1481	TRUE	0.8519	TRUE	C	T
7840	0.5	0.9221	TRUE	0.4610	TRUE	0.5390	TRUE	T	C
7840	1	0.7817	TRUE	0.6091	TRUE	0.3909	TRUE	T	C
7840	1.5	0.6494	TRUE	0.6753	TRUE	0.3247	TRUE	T	C
8334	0.5	0.0004	FALSE	0.9998	TRUE	0.0002	FALSE	T	C
8334	1	0.0031	FALSE	0.9984	TRUE	0.0016	FALSE	T	C
8334	1.5	0.0001	FALSE	1.0000	TRUE	0.0000	FALSE	T	C
8413	0.5	0.6872	TRUE	0.3436	TRUE	0.6564	TRUE	T	C
8413	1	0.0172	FALSE	0.9914	TRUE	0.0086	FALSE	T	C
8413	1.5	0.0050	FALSE	0.9975	TRUE	0.0025	FALSE	T	C
8968	0.5	0.0000	FALSE	0.0000	FALSE	1.0000	TRUE	C	T
8968	1	0.1574	TRUE	0.0787	TRUE	0.9213	TRUE	C	T
8968	1.5	0.0436	FALSE	0.0218	FALSE	0.9782	TRUE	C	T
9321	0.5	0.2847	TRUE	0.1424	TRUE	0.8576	TRUE	C	T
9321	1	0.3130	TRUE	0.1565	TRUE	0.8435	TRUE	C	T
9321	1.5	0.1416	TRUE	0.9292	TRUE	0.0708	TRUE	C	T
Total			19		10		14		

Appendix C: Word Recall List

Table 13: Word Recall Lists

A	B	C	D	E	F
crevice	bastard	streak	bastard	league	crevice
motion	boycott	faucet	boycott	hippie	kernel
transit	cluster	impulse	cluster	ascent	clement
goblet	immense	assure	immense	tension	section
source	calcium	auction	calcium	glamour	prolong
gender	texture	vandal	texture	dismiss	source
inspect	outcome	fodder	outcome	vacant	concoct
outlook	affect	saloon	affect	devise	inkling
awkward	mooring	direct	mooring	passion	convict
worship	clobber	breast	clobber	graphic	invest
resort	startle	supreme	startle	ghetto	sermon
leaflet	jargon	dismiss	jargon	strung	avenge
request	region	ruffle	region	breach	locate
treble	trinket	racist	trinket	sphere	orchard
sector	poacher	liquor	poacher	access	ensure
lavish	symbol	hazard	symbol	diverse	platter
random	anthem	closure	anthem	immense	treble
syringe	payroll	release	payroll	outcome	archer
regime	absurd	divert	absurd	affect	adverse
burner	senate	climax	senate	absurd	lavish
poultry	swindle	mingle	swindle	swindle	persist
cashier	proper	deviant	proper	captive	prevent
legion	resign	percent	resign	anxious	amongst
accuse	escort	strode	escort	dresser	oblige
campus	adjust	vacant	adjust	billion	burner
tandem	precise	devise	precise	quibble	revoke
grenade	sultan	passion	sultan	comment	coward
patrol	premium	pallet	premium	concept	memoir
mission	former	graphic	former	mighty	corpse
suburb	captive	expense	captive	vaccine	output
G	H	I	J	K	L
streak	conform	crevice	shudder	posture	conform
faucet	cockney	symptom	hippie	tumbler	script

impulse	sought	turban	hammock	nipple	estate
assure	matron	bronze	ascent	gender	current
auction	wigwam	clergy	morsel	stance	faucet
vandal	ascent	gender	employ	paisley	auction
fodder	whimper	cheque	crisis	fatigue	supreme
saloon	galley	stingy	current	nimble	release
direct	streak	inspect	impulse	govern	ghetto
breast	vandal	stance	assure	mustang	credit
supreme	saloon	convict	dismiss	request	access
dismiss	direct	caddie	ruffle	treason	contend
ruffle	hazard	nimble	hazard	propose	meagre
racist	closure	sermon	closure	relish	bomber
liquor	mingle	tribute	mingle	lavish	diverse
hazard	vacant	pursue	percent	persist	texture
closure	sulphur	butler	devise	parlor	outcome
release	county	intense	passion	syringe	startle
divert	boycott	preview	bomber	amongst	absurd
climax	cluster	convent	boycott	regime	escort
mingle	poacher	adverse	immense	critic	adjust
deviant	swindle	madness	texture	glacial	premium
percent	captive	culture	startle	shrewd	captive
strode	facial	neutral	poacher	shroud	drought
vacant	midget	syringe	symbol	exclaim	impact
devise	impact	oblige	senate	trench	quibble
passion	ferret	revoke	facial	cashier	pulpit
pallet	passive	trench	parish	restore	blunder
graphic	concept	conceal	concept	accuse	mighty
expense	pulpit	restore	unique	tandem	picket

Appendix D: Pre and Post-Experiment Questionnaires

Clear All Button

Show Hide

CLEAR ALL

Small Motor Movement Cognitive Effects

ID:

Pre-screening Questionnaires

What hand are you?

- Left
- Right
- Both

For the following please indicate which hand you primarily use to...

- | | | |
|--------------------------------|--------------------------------|--------------------------------|
| Write? | Throw a ball? | Operate a computer mouse? |
| <input type="checkbox"/> Left | <input type="checkbox"/> Left | <input type="checkbox"/> Left |
| <input type="checkbox"/> Right | <input type="checkbox"/> Right | <input type="checkbox"/> Right |

How would you characterize your sleep last night?

- Very Poor
- Poor
- Fair
- Good
- Very Good

Did you consume any products with caffeine today?

- Yes
- No

If yes:

What product did you consume?

When did you last consume this product?

Approximately how many ounces of this product have you consumed today?

Did you exercise today?

- Yes
- No

If yes:

Did you finish your exercise at least 2 hours ago?

- Yes
- No

What type of exercise did you engage in?

How long did you engage in the exercise (in minutes)?

"Clear All" Button

Show Hide

CLEAR ALL

Small Motor Movement Cognitive Effects

ID:

Post-Experiment Questionnaire

(Page 1 of 2)

Age?

Are you Male or Female?

Male

Female

Prefer not to answer

What's your highest education level?

Video Game Experience:

Do you play video games? This includes all types of video games (action, word, puzzle, strategy, etc.) played on any platform (personal computer, tablet, smartphone, console, etc.)

Yes

No

If yes:

On average, how many hours per week do you play video games?

How many years have you been playing video games?

Please list the names of the video games you play most frequently?

What platform(s) do you play video games on? (Check all that apply)

Personal laptop

Console

Tablet

Smartphone

Handheld device

Other, please list below

Exercise Habits:

On average, how many days per week do you exercise?








On average, how many minutes long in your typical exercise session?

What type of exercise do you engage in most frequently?








Continued on page 2...

Small Motor Movement Cognitive Effects
Post-Experiment Questionnaire
 (Page 2 of 2)

Rate the '**Difficulty**' of each task below:

		Very Low	Low	Little Low	Little High	High	Very High
	'x'Box	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	OddBall	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Go-NoGo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Word Recall	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Dot Matrix	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Mouse & Key	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Stare	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
		Very Low	Low	Little Low	Little High	High	Very High

Rate your '**Performance**' on each task below:

		Very Low	Low	Little Low	Little High	High	Very High
	'x'Box	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	OddBall	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Go-NoGo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Word Recall	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Dot Matrix	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Mouse & Key	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	Stare	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
		Very Low	Low	Little Low	Little High	High	Very High

If you have any comments, please type them in here:

Appendix E: Word Recall Time Score Pseudo Code

```
Build new table only containing participant observations with word Matches
For each observation {
    Set 'm' = 15, 'w' = 6.6
    Use word entry time info as 'x' for following equation:  $\exp(-x/m)*w$  and enter
    result into new column labeled time_score
}
Create new table called Word_Summary
For each participant {
    For each condition {
        Sum all word scores and time scores and add as one row in
        Word_Summary table
    }
}
Run t-test's and other desired analysis on Word_Summary table results
```

Appendix F: Approval Letter



DEPARTMENT OF THE AIR FORCE
AIR FORCE RESEARCH LABORATORY
WRIGHT-PATTERSON AIR FORCE BASE OHIO 45433

AUG 25 2016

MEMORANDUM FOR AFIT/ENG (BRETT BORGHETTI)

FROM: 711 HPW/IR

SUBJECT: IRB Approval for the Use of Human Volunteers in Research

1. Protocol title: Small Motor Movement Cognitive Effects
2. Protocol number: FWR20160127 H
3. Protocol version: v1.00
4. Risk: Minimal
5. Approval date: 25 August 2016
6. Expiration date: 24 August 2017
Your renewal submission date is *one month prior* to your expiration date. The renewal is due 24 July 2017.
7. Expedited Review Category:
 32CFR219.110(b)(1) 32CFR219.110(b)(2) 32CFR219.110(b)(3)
 32CFR219.110(b)(4) 32CFR219.110(b)(5) 32CFR219.110(b)(6)
 32CFR219.110(b)(7) 32CFR219.109: Full Board
8. Assurances and Agreements with Expiration Dates:
 - a. AFRL DoD Assurance 50002: 6 March 2017
 - b. Wright State University Assurance FWA00002427: 10 April 2017
9. The study objective is to determine the effect of small motor activity in the hand/fingers (such as mouse and keyboard movement) on cognitive abilities. Since the effects of hand / finger movements on cognitive performance are largely unknown, experiments will be conducted in two phases. In the proposed study we will screen a large number of potential cognitive skills tasks over a small number of participants to determine which areas show evidence of being affected by hand/finger movements. Assuming cognitive effects are observed in this study, a future study will investigate if the same set of cognitive improvements persists within a larger group of participants while collecting EEG and other physiological signals.

10. All inquiries and correspondence concerning this protocol should include the protocol number and name of the primary investigator. Please contact the 711 HPW/IR office using the organizational mailbox at AFRL.IR.ProtocolManagement@us.af.mil or calling 937-904-8100 [DSN 674].



KIM E. LONDON, JD, MPH, CIP
Chair, AFRL IRB

1st Indorsement to AFIT/ENG (BRETT BORGHETTI), 25 August 2016, Approval for Use of Humans in Research, Initial Review - Expedited, FWR20160109H

MEMORANDUM FOR AFMSA/SGE-C

This protocol has been reviewed and approved by the AFRL IRB. I concur with the recommendation of the IRB and approve this research.



TIMOTHY J. SAKULICH
Vice Director
711th Human Performance Wing

Bibliography

- [1] R. J. Murnane, J. B. Willett, and F. Levy, "The Growing Importance of Cognitive Skills in Wage Determination," *Rev. Econ. Stat.*, vol. 77, no. 2, pp. 251–266, 1995.
- [2] Y. K. Chang, J. D. Labban, J. I. Gapin, and J. L. Etnier, "The effects of acute exercise on cognitive performance: A meta-analysis," *Brain Research*, vol. 1453, pp. 87–101, 2012.
- [3] T. A. Plerhoples, Y. Zak, T. Hernandez-Boussard, and J. Lau, "Another use of the mobile device: Warm-up for laparoscopic surgery," *J. Surg. Res.*, vol. 170, no. 2, pp. 185–188, 2011.
- [4] J. C. Rosser Jr, D. a Gentile, K. Hanigan, and O. K. Danner, "The effect of video game 'Warm-up' on performance of laparoscopic surgery tasks," *J. Soc. Laparoendosc. Surg.*, vol. 16, no. 713, pp. 3–9, 2012.
- [5] J. Y. Lee, P. Mucksavage, D. C. Kerbl, K. E. Osann, H. N. Winfield, K. Kahol, and E. M. McDougall, "Laparoscopic warm-up exercises improve performance of senior-level trainees during laparoscopic renal surgery," *J. Endourol.*, vol. 26, pp. 545–550, 2012.
- [6] T. S. Lendvay, T. C. Brand, L. White, T. Kowalewski, S. Jonnadula, L. D. Mercer, D. Khorsand, J. Andros, B. Hannaford, and R. M. Satava, "Virtual reality robotic surgery warm-up improves task performance in a dry laboratory environment: A prospective randomized controlled study," *J. Am. Coll. Surg.*, vol. 216, no. 6, pp. 1181–1192, 2013.
- [7] J. C. Rosser Jr, P. J. Lynch, L. Cuddihy, D. A. Gentile, J. Klonsky, and R. Merrell, "The Impact of Video Games on Training Surgeons in the 21st Century," *Arch. Surg.*, vol. 142, no. 2, pp. 181–186, 2007.
- [8] LegacyAccount, "How do you warm up your fingers before a game? | IGN Boards," *IGN Boards*, 2009. [Online]. Available: <http://www.ign.com/boards/threads/how-do-you-warm-up-your-fingers-before-a-game.177189646/>. [Accessed: 02-Jun-2016].
- [9] A. P. Tvaryanas, W. T. Thompson, and S. H. Constable, "Human factors in remotely piloted aircraft operations: HFACS analysis of 221 mishaps over 10 years," *Aviat. Sp. Environ. Med.*, vol. 77, no. 7, pp. 724–732, 2006.
- [10] S. V. Onyper, T. L. Carr, J. S. Farrar, and B. R. Floyd, "Cognitive advantages of chewing gum. Now you see them, now you don't," *Appetite*, vol. 57, no. 2, pp. 321–328, 2011.
- [11] "Cognition Laboratory Experiments." [Online]. Available: <http://psych.hanover.edu/JavaTest/CLE/Cognition/Cognition.html>.
- [12] J. L. J. L. Otto and B. J. Webber, "Mental health diagnoses and counseling among pilots of remotely piloted aircraft in the United States Air Force," *MSMR*, vol. 20, no. 3, pp. 3–8, 2013.
- [13] P. D. Tomporowski, "Effects of acute bouts of exercise on cognition," *Acta Psychol. (Amst.)*, vol. 112, no. 3, pp. 297–324, 2003.
- [14] M. B. Jalink, E. Heineman, J. P. E. N. Pierie, and H. O. ten Cate Hoedemaker,

- “The effect of a preoperative warm-up with a custom-made Nintendo video game on the performance of laparoscopic surgeons,” *Surg. Endosc. Other Interv. Tech.*, vol. 29, no. 8, pp. 2284–2290, 2015.
- [15] S. V. Onyper, T. L. Carr, J. S. Farrar, and B. R. Floyd, “Cognitive advantages of chewing gum. Now you see them, now you don’t,” *Appetite*, vol. 57, no. 2, pp. 321–328, 2011.
- [16] P. S. Cooper, Á. Darriba, F. Karayanidis, and F. Barceló, “Contextually sensitive power changes across multiple frequency bands underpin cognitive control,” *Neuroimage*, vol. 132, pp. 499–511, 2016.
- [17] E. Hogervorst, W. Riedel, a Jeukendrup, and J. Jolles, “Cognitive performance after strenuous physical exercise.,” *Percept. Mot. Skills*, vol. 83, no. 1971, pp. 479–488, 1996.
- [18] A. P. Allen, T. J. C. Jacob, and A. P. Smith, “Effects and after-effects of chewing gum on vigilance, heart rate, EEG and mood,” *Physiol. Behav.*, vol. 133, pp. 244–251, 2014.
- [19] K. Sakamoto, H. Nakata, and R. Kakigi, “The effect of mastication on human cognitive processing: A study using event-related potentials,” *Clin. Neurophysiol.*, vol. 120, no. 1, pp. 41–50, 2009.
- [20] J. C. Rosser Jr, P. J. Lynch, L. Cuddihy, D. A. Gentile, J. Klonsky, and R. Merrell, “The Impact of Video Games on Training Surgeons in the 21st Century,” *Arch. Surg.*, vol. 142, no. 2, pp. 181–186, 2007.
- [21] E. Tulving, “What is episodic memory?,” *Curr. Dir. Psychol. Sci.*, vol. 2, no. 3, pp. 67–70, 1993.
- [22] P. S. Cooper, Á. Darriba, F. Karayanidis, and F. Barceló, “Contextually sensitive power changes across multiple frequency bands underpin cognitive control,” *Neuroimage*, vol. 132, pp. 499–511, 2016.
- [23] E. Bronnimann, H. Hoffmann, J. Schafer, D. Hahnloser, and R. Rosenthal, “Effect of different warm-up strategies on simulated laparoscopy performance: A randomized controlled trial,” *J. Surg. Educ.*, vol. 72, no. 1, pp. 96–103, 2015.
- [24] J. C. Rosser Jr, D. a Gentile, K. Hanigan, and O. K. Danner, “The effect of video game ‘Warm-up’ on performance of laparoscopic surgery tasks,” *J. Soc. Laparoendosc. Surg.*, vol. 16, no. 713, pp. 3–9, 2012.
- [25] A. P. Allen, T. J. C. Jacob, and A. P. Smith, “Effects and after-effects of chewing gum on vigilance, heart rate, EEG and mood,” *Physiol. Behav.*, vol. 133, pp. 244–251, 2014.
- [26] V. B. Tupov, “ON CLERICAL TASKS ASSOCIATED WITH ADMINISTRATION OF PEPPERMINT ODOR,” no. July, pp. 12–16, 2015.
- [27] H. AARTS and A. DIJKSTERHUIS, “The Automatic Activation of Goal-Directed Behaviour: the Case of Travel Habit,” *J. Environ. Psychol.*, vol. 20, no. 1, pp. 75–82, 2000.
- [28] E. Tulving and W. Donaldson, “Episodic and Semantic Memory,” *Acad. Press. Inc.*, no. Organization of Memory, 1972.
- [29] A. Baddeley, “The epissodic buffer: a new component of working memory?,” *Trends Cogn. Sci.*, vol. 4, no. 11, pp. 417–423, 2000.

- [30] A. P. Allen and A. P. Smith, “Effects of chewing gum and time-on-task on alertness and attention.” *Nutr. Neurosci.*, vol. 15, no. 4, pp. 176–85, 2012.
- [31] A. P. Smith, “Effects of chewing gum on cognitive function, mood and physiology in stressed and non-stressed volunteers,” *Nutr. Neurosci.*, vol. 13, no. 1, pp. 7–16, 2010.
- [32] “Home — PsychoPy v1.85.0.” [Online]. Available: <http://www.psychopy.org/>. [Accessed: 27-Feb-2017].
- [33] I. J. Deary, D. Liewald, and J. Nissan, “A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task.” *Behav. Res. Methods*, vol. 43, no. 1, pp. 258–68, 2011.
- [34] A. Jain, R. Bansal, A. Kumar, and K. D. Singh, “A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students.” *Int. J. Appl. basic Med. Res.*, vol. 5, no. 2, pp. 124–7, 2015.
- [35] “Human Benchmark - Reaction Time Statistics.” [Online]. Available: <http://www.humanbenchmark.com/tests/reactiontime/statistics>.
- [36] M. Mendl, “Performing under pressure: Stress and cognitive function,” *Appl. Anim. Behav. Sci.*, vol. 65, no. 3, pp. 221–244, 1999.

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> OMB No. 074-0188		
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 23-03-2017		2. REPORT TYPE Master's Thesis		3. DATES COVERED (From - To) Aug 2015 - March 2017	
TITLE AND SUBTITLE Small-muscle Movement and Its Effect on Cognitive Functions			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
AUTHOR(S) Meyer, Timothy R., Capt			5d. PROJECT NUMBER 16G347		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENG-MS-17-M-053		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFIT/ENR 2950 Hobson Way WPAFB OH 45433-7765			10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR (managed by AFIT Faculty Research Council)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.					
13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.					
14. ABSTRACT This thesis study provides the ground work for exploring ways to improve the cognitive state of an operator as needed. Small-muscle warm-up such as hand and finger movement, is widely unexplored for its effects on cognitive performance. The experiment investigates the use of a hand and finger warm-up activity and its effects on reaction time, categorization, impulse control, episodic memory, and visual spatial (VS) working memory. Sixteen participants each completed five cognitive measurement activities under two different conditions: preceded by a hand and finger warm-up activity and also preceded by a relaxed state activity. Results showed that hand and finger activity significantly increased reaction time in three measures of cognitive performance, the largest increase being a reaction time of 8 milliseconds faster in the treatment condition than under the control condition. The analysis also suggested an increase in episodic memory. In addition to the effect of the warm-up, the pace with which participants conducted the warm-up has a significantly positive correlation coefficient of 0.54 with regards to the size of the cognitive effect for reaction time. The findings of this study show that small muscular movements do have a significant effect on cognitive functions.					
15. SUBJECT TERMS Small-Muscle, Cognitive Functions, Performance, Warm-up, Exercise, Hand & Finger Movement, Reaction Time, Categorization, Impulse Control, Working Memory, Episodic Memory					
16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 117	19a. NAME OF RESPONSIBLE PERSON Dr. Brett Borghetti, AFIT/ENG
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U			19b. TELEPHONE NUMBER (Include area code) 937-255-3636, ext. 4612 (NOT DSN) brett.borghetti@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18