GENDER AND PERSONALITY DIFFERENCES IN COGNITIVE TASKS: A PERFORMANCE AND MENTAL WORKLOAD STUDY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BETÜL BATUN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INDUSTRIAL ENGINEERING

JUNE 2015

Approval of the thesis:

GENDER AND PERSONALITY DIFFERENCES IN COGNITIVE TASKS: A PERFORMANCE AND MENTAL WORKLOAD STUDY

submitted by **BETÜL BATUN** in partial fulfillment of the requirements for the degree of **Master of Science in Industrial Engineering Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver ____________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Murat Köksalan ____________________
Head of Department, **Industrial Engineering**

Prof. Dr. Canan Çilingir ____________________
Supervisor, **Industrial Engineering Dept., METU**

Assist. Prof. Dr. Murat Perit Çakır ____________________
Co-Supervisor, **Cognitive Science Dept., METU**

Examining Committee Members:

Assoc. Dr. Canan Sepil ____________________
Industrial Engineering Dept., METU

Prof. Dr. Canan Çilingir ____________________
Industrial Engineering Dept., METU

Assist. Prof. Dr. Murat Perit Çakır ____________________
Cognitive Science Dept., METU

Assoc. Dr. Sedef Meral ____________________
Industrial Engineering Dept., METU

Assoc. Dr. Ferda Can Çetinkaya ____________________
Industrial Engineering Dept., Çankaya University

Date: 24/06/2015

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Betül Batun

Signature:

ABSTRACT

GENDER AND PERSONALITY DIFFERENCES IN COGNITIVE TASKS:
A PERFORMANCE AND HUMAN MENTAL WORKLOAD STUDY

Batun, Betül
M.S., Department of Industrial Engineer
Supervisor: Prof. Dr. Canan Çilingir
Co-Supervisor: Assist. Prof. Dr. Murat Perit Çakır

June 2015, 154 Pages

The increasing complexity of today's systems creates more demand on individuals during their daily tasks and mental workload becomes more important during a cognitive process. In this study, our aim is to investigate whether individual differences, cognitive task types and task difficulty have a significant impact on mental workload and how those elements affect mental workload if they do. N-back, maze and information sampling task (IST) are used as cognitive tasks. Mental workload is assessed by monitoring changes in blood oxygenation using functional near-infrared spectroscopy (fNIR). It is observed that males do more trials than females during IST, males have better 4-back performance but no gender differences are observed for other performance measures. No significant difference between genders exists in blood oxygenation levels. Agreeableness and conscientiousness personality traits have negative correlations with IST performance. Extraversion, agreeableness, conscientiousness and openness to experience have correlations with blood oxygenation levels. Lastly, we observe from our results that performance decreases with the increasing difficulty level and oxygenation level in voxels increases with the increasing difficulty level. In the future, a larger sample of participants, more difficult tasks and participants from different fields and schools should be chosen in order to provide more reliable results.

Keywords: Cognitive task, mental workload, gender differences, personality traits, fNIR

ÖZ

BİLİŞSEL GÖREVLERDE CİNSİYET VE KİŞİLİK ÖZELLİKLERİ FARKLILIKLARI: PERFORMANS VE ZİHİNSEL İŞ YÜKÜ ÇALIŞMASI

Batun, Betül
Yüksek Lisans, Endüstri Mühendisliği
Tez Yöneticisi: Prof. Dr. Canan Çilingir
Ortak Tez Yöneticisi: Yrd. Doç. Dr. Murat Perit Çakır

Haziran 2015, 154 Sayfa

Günümüz sistemlerinin karmaşıklığı günlük görevleri sırasında bireylerden daha fazla yararlanmaya yönelik bir sistem yaratmış ve bu da karşımıza dinamik bir yapıya sahip olan zihinsel iş yükünü çıkarmıştır. Bu çalışmadaki hedefimiz, cinsiyetin ve kişilik faktörlerinin, bilişsel görev tipinin, görev zorluk derecesinin performansa ve zihinsel iş yüküne olan etkisini araştırmaktır. N-geri, labirent ve bilgi seçme görevleri bu çalışmada bilişsel görevler olarak kullanılmıştır. Zihinsel iş yükü ölçümü için fonksiyonel yakın kızılötesi spektroskopi (fNIR) cihazı kullanılmıştır. Sonuçlara göre erkekler bilgi seçme görevinde kadınlardan daha çok deneme yapmışlar ve erkekler 4-geri görevinde kadınlardan daha iyi bir performans göstemişlerdir. Diğer görevlerdeki performanslarda ise cinsiyet farkına rastlanmamıştır. Ayrıca kanlanma değerleri arasında farklılık olmaması, bilişsel görevlerde kadın ve erkeğin zihinsel yorgunlukları arasında bir fark olmadığını göstermektedir. Kişilik faktörlerinin performansla ilişkisine bakıldığında, uyumluluk ve sorumluluk faktörleri ile bilgi seçme performans ölçümleri arasında negatif bir korelasyon gözlemlenmiştir. Dışadönüklük, uyumluluk, sorumluluk ve deneyime açık olma faktörlerinin kanlanma değerleriyle korelasyonları vardır. Bir başka amaç da performansı ve kanlanma seviyelerini farklı zorluk derecelerinde incelemektir. Sonuçlara göre, beklenildiği gibi, zorluk seviyesi arttıkça kişilerin doğru yapma seviyeleri azalmakta, kanlanma seviyeleri artmaktadır. İleriki çalışmalarda daha geniş bir örneklem, daha zor bilişsel görevler, daha farklı okullardan ve daha farklı bölümlerden katılımcıların kullanılması daha güvenilir sonuçlar sağlayacaktır.

Anahtar kelimeler: Bilişsel görevler, zihinsel iş yükü, cinsiyet farkı, kişilik farkı, Fnir

To my grandmother

ACKNOWLEDGEMENTS

First and foremost, I would like to express the deepest appreciation to my supervisor Prof. Dr. Canan Çilingir for the valuable guidance, moral support and friendly tolerance throughout this study. She inspired me greatly to work in this study and she was much more than an advisor to me. I cannot thank her enough for this great experience.

I would like to give special thanks to Assist. Prof. Dr. Murat Perit Çakır, my co-supervisor for his valuable supervision and insightful comments in every step of the study. I am grateful to him for his guidance and positive attitude.

I would also thank to Assoc. Dr. Canan Sepil, Assoc. Dr. Sedef Meral and Assoc. Dr. Ferda Can Çetinkaya for them being in my examining committee and their valuable contributions to this study.

I cannot thank my sister Sakine Batun enough. Without her support, understanding, love and friendship this work could not be completed. Her contribution means more than I could express just by words.

I am beyond grateful to my parents, Ahmet Batun and Ayşe Batun who supported me along the way, making this effort possible. For the last three years, I have been devoted to this thesis study, which has been hard work but finally paid off. I am grateful to my family, my friends and all those who accompanied and supported me during this journey.

Last but not least, I would like to thank all participants in the experiments for their time and contribution.

TABLE OF CONTENTS

FIGURES

TABLES

CHAPTER 1

INTRODUCTION

Recent advances and the increasing complexity of today's technology create more demand on individuals while performing everyday tasks. This increasing demand becomes more difficult to be satisfied and brings an excessive burden over limited information processing capacity of a person. Especially in complex systems based on human control in time-sensitive contexts, the human factor should be considered as an important factor. For example, air traffic control (ATC) to maintain traffic safety and efficiency, and human operations in health care are examples of high hazard industries that need to be considered from a cognitive processing capacity perspective.

Mental workload is the operator's excess capacity between demands of the task and his/her current capacity limits while achieving adequate task performance (Jex, 1988). Human mental workload is encountered in many complex systems. For example: Air traffic controllers experience mental workload while remembering positions and heights of airplanes from the ground level. Any complex or long surgery forces a doctor's mental strain. Long hours of meeting in management based systems can generate high levels of cognitive workload. Examples of human mental workload are easy to observe in any complex system.

Human mental workload, which is the margin between the current task demand and the operator's motivated capacity, has a dynamic structure (Jex, 1988). The complexity of a task may change according to the task type, the capability of performing a task may differ depending on the variety of people and both have impacts on human mental workload. Individual differences, cognitive task types and task difficulty are thought as they have an effect on mental workload. However, whether these factors have really an effect on mental workload and how these elements affect mental workload if they do, are open questions waiting to be answered. This is the main motivation behind our study.

Functional near infrared spectroscopy (fNIR) is used within the study as a physiological method to assess mental workload. Studies show that human brain hemodynamic activity assessed by fNIR can provide a sensitive and reliable assessment of human mental workload during a complex task (Chance *et* al., 1998; Izzetoglu *et* al. 2004; Ayaz *et* al., 2008). In this study, we used fNIR as a measure tool of mental workload and tried to explain relationships between mental workload and individual differences of users.

1.1. The Purpose of the Study

The aim of this study is to investigate the effects of gender on different cognitive tasks by measuring performance and monitoring changes in blood oxygenation using fNIR to assess mental workload. Another purpose is to find out if there are any personality differences on performances and blood oxygenation levels during cognitive tasks. The findings for these aims will help literature about how individual differences affect mental workload by investigating both gender and personality differences.

This study investigates whether performance and fNIR measures change according to the task difficulty level. Increase in the difficulty of the task may cause reduction in performance and rising activation in frontal cortical areas. Difficulty degrees of different cognitive tasks are explored by fNIR measures. Since fNIR has been used to assess mental workload, various types of cognitive tasks may cause different levels of fatigue according to their difficulty. The results will provide information about comparison of different kinds of cognitive tasks. In the literature, comparing both levels of cognitive tasks and different task types does not exist. Therefore, this study will contribute.

1.2. Research Questions

Research questions of this study are determined according to the purpose of the study. The following research questions are investigated through experiments conducted in this study:

- Do performance levels of subjects change according to gender, when the same cognitive tasks are performed?
- Do performance levels of subjects change according to gender at different levels of the same cognitive tasks?
- Do performance levels of subjects change according to the difficulty levels of the same cognitive tasks?
- Do changes in blood oxygenation levels of subjects vary according to gender, when the same cognitive tasks are performed?
- Do changes in blood oxygenation levels of subjects vary according to gender at different levels of the same cognitive tasks?
- Do changes in blood oxygenation levels of subjects change according to the difficulty levels of the same cognitive tasks?
- Is there a difference between changes in oxygenation levels of left dorso-lateral prefrontal cortex (DLPFC) and right DLPFC, during a cognitive task?
- Do changes in blood oxygenation levels of subjects change according to type of the cognitive tasks?

- Is there a relationship between changes in blood oxygenation and performance level during a cognitive task?
- Do performance levels of subjects change according to personality traits, when the same cognitive tasks are performed?
- Do performance levels of subjects change according to personality traits at different levels of the same cognitive tasks?
- Do changes in blood oxygenation of subjects change according to personality traits, when the same cognitive tasks are performed?
- Do changes in blood oxygenation of subjects change according to personality traits at different levels of the same cognitive tasks?

1.3. Organization of the Thesis

This chapter explains the aim of the study and presents the research questions. Chapter 2 presents a summary of the literature related to this study. Chapter 3 gives information about the experimental procedure including the participants, the apparatus, the experimental process and the cognitive tasks used. Chapter 4 presents the results of the statistical analysis and provides a general discussion of the findings. Chapter 5 includes conclusions, limitations and recommendations.

CHAPTER 2

LITERATURE REVIEW

In this chapter, mental workload and its measure techniques will be briefly reviewed. Then, types of cognitive tasks, gender differences and lastly personality differences during cognitive tasks will be explained.

2.1. Mental Workload

A generally accepted definition of mental workload does not exist, because it covers a broad range of factors. Jex (1988, p.9) defines mental workload as "… the operator's evaluation of the attentional load margin (between their motivated capacity and the current task demands) while achieving adequate task performance in a mission-relevant context." De Waard (1996) offers an alternative definition of mental workload by saying that it is the specification of the amount of information processing capacity that is used for task performance. Gopher & Braune (1984) stated that mental workload can be defined as a mental strain resulting from performing a task. It is a concept to explain the inability of human operators to cope with the requirements of a task. Mental workload can vary by task complexity, motivation and fatigue level of the operator. Therefore, it can be said that mental workload is both task- and person-specific. These factors change mental workload, but do not change its definition.

Measuring mental workload may provide insights about the design of the system, the possible sources of human errors, and the reasons underlying environmental problems such as noise factors, luminance. Such insights can be used to design better systems. Three main categories of mental workload measure are defined in the literature: subjective measures, performance measures, and physiological measures (Eggemeier, 1988; De Waard, 1996). Each measure is defined and explained under its own title in the following paragraphs.

2.1.1. Subjective Ratings

In this technique, the participant either answers a questionnaire or rates his/her perceived mental workload in a specific situation. If subjective ratings are well-structured, they may be helpful in learning the situation of the participant directly. There are many subjective assessment methods. The frequently used ones are as follows: Bedford scales (Wierwille and Casali, 1983), Overall Workload (Vidulich and Tsang, 1987), NASA TLX (Hart and Staveland, 1988), SWAT (Reid and Nygren, 1988), and W/Index (Workload Index: North and Riley, 1989). Although they are cheap and easy to use, the repeatability and validity of these instruments can be uncertain. To get around this problem, other methods are often used together with subjective ratings.

2.1.2. Performance Measures

Performance measures of workload can be grouped into two types: primary task measures and secondary task measures. Primary task measure is based on measuring the performance observed during the primary task, but the performance level alone is not considered as an indicator of mental effort because of strategic reallocation of mental capacity (Wilson, 2004). Thus, simply considering performance to assess the participant’s mental effort may not be sufficient. There are many primary task measures and the most frequently used ones are as follows: Accuracy, number of errors, reaction times, and speed (O’Donnell & Eggemeier, 1986).

Secondary task measure is based on measuring the performance of a secondary task while the participant is performing a primary task. For example, performing basic arithmetic calculations during driving and talking to other pilots during flight are secondary tasks for a driver and a pilot, respectively. The secondary task measure is classified into two methodologies: Auxiliary Task and Loading Task (O'Donnell and Eggemeier, 1994). Secondary task measure is less frequently practiced as compared to subjective ratings. Colle and Reid (1999) claimed that globally sensitive secondary task measure methods are needed. However, different primary tasks require different secondary tasks, which create generalizability issues (Meshkati & Lowewinthal, 1988).

2.1.3. Physiological Measures

Physiological measure is an objective measure approach based on physical reactions of the participant's body as monitored by specific sensors. These physical reactions can be monitored in terms of skin conductivity, cardiovascular activity, respiratory responses, brain activity or pupillary size, each of which are measured by a dedicated equipment specialized on that modality (Wilson & O'Donnell, 1988).

Heart rate (HR) and heart rate variability (HRV) are the most commonly used measures of workload, because the heart beat is a relatively easy measure to obtain (Wilson and O'Donnell' 1988). Generally, increased HR is associated with increased workload (Jorna, 1993; Roscoe, 1993; Wilson, Fullenkamp & Davis, 1994; Hankins & Wilson, 1998), but some studies argued that there are other psychological, physical and environmental factors that could affect HR (Lee & Parks, 1990; Roscoe, 1992). Therefore, to measure mental workload with increased/decreased HR could be misleading.

Some studies showed that heart rate variability (HRV) decreases under high workload levels (Mulder, 1979; Vincente, Thornton & Moray, 1987; Mulder & Mulder, 1997), but there are also studies that reported conflicting results where increased workload did not lead to a decrease in HRV (Brookings, Wilson, & Swain, 1996; Hankins & Wilson, 1998). Therefore HRV is not a totally reliable measure for cognitive workload. Like HR, HRV may be affected by not only mental workload but also external factors.

The electroencephalogram (EEG) is a workload measure that uses surface electrodes placed on a participant's scalp to monitor the changes in electric potentials due to the brain's activity during a specific task. (DE Waard, 1996). EEG focuses on brain activity of a person. There are some types of bands (delta, theta, alpha, beta and ultra beta) that are used with this method to measure changes in the electrical potential in the brain cells (Brookings *et* al., 1996). Simulation based studies (Borghini *et* al., 2011; Dussault *et* al., 2005) showed that there is a power increase in theta band and changes in alpha band with low performance. Electrons on front-middle and top-middle also showed a power increase in theta band with distraction due to task difficulty. However, physical movements of the subjects can be a problem with this method. Physical limitations, non-practical structure and the cost of implementation of EEG can cause a problem during application.

Electrooculogram (EOG) measures the velocity of the saccadic eye movements to find workload (Galley, 1993). Borgini *et* al. (2014) used EOG in their study. Their results showed when there was an increase in the workload of the pilot; increases in frequency and length of blinking, a reduction in focusing were observed. However, there are external factors such as the amount of light in the cockpit that can also change the pupil size and blink frequencies (De Rivecourt *et* al., 2008). In short, EOG offers good indicator for visual workload, but it is not enough to see total cognitive workload.

Positron Emission Tomography (PET) is an important imaging technique for measuring blood flow or energy consumption in the brain. PET is an invasive technique that requires the injection of harmless radioactive tracers in the blood stream. The half-life of the tracer limits the PET experiments to relatively small durations, and the subjects are required to stay in a confined position to ensure accurate detection of brain activity. Single Photon Emission Computed Tomography (SPECT) is another brain imaging technique which is similar to PET, but it involves less detail and is less expensive than PET (Greely & Wagner, 2011). The image quality in SPECT makes functional brain analysis challenging.

Functional magnetic resonance imaging (fMRI) is another popular and common neuroimaging technique. It allows detecting increases in blood oxygen levels and shows detailed maps of the brain areas underlying human mental activities (Greely & Wagner, 2011). It is very useful to detect complex emotions in the brain, but it is costly and hard-to-use as a neuroimaging tool (Cansiz, 2012). Moreover, it requires subjects to be monitored in an MRI scanner in a confined position, which makes it difficult to conduct practical studies in human factors research.

Functional near-infrared spectroscopy (fNIR) is a relatively new neuroimaging technology compared to commonly employed techniques such as EEG, PET, SPECT and fMRI. It has been used in functional brain studies to monitor changes in the concentration of oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HbR).

After Chance's (1998) pioneering work on the development of fNIR technology and his demonstration of its use for monitoring oxygenation changes in the brain due to the modulation of specific cognitive processes, many follow-up studies consistently produced compatible results for the monitoring of mental workload with fNIR. Izzetoglu *et* al. (2004) used fNIR to help to understand the cognitive state of a user during a complex task. The complex task was a video game called the Warship Commander Task (WCT). They used different task difficulty and task load levels by changing the state of the game. The main aim of the study was to see that fNIR was an appropriate measure to predict changes in cognitive workload. Therefore, the hypothesis of the study was that blood oxygenation in the prefrontal cortex which was assessed by fNIR, would have a raise with increasing task difficulty and have a positive correlation with performance. When a task became too difficult, the participant was expected to disengage his/her attention on the cognitive task and the blood oxygenation level was expected to decrease. Results showed that change in the blood oxygenation was significantly sensitive to task difficulty and had a positive correlation with performance. So, it indicated that fNIR could be reliably used to monitor hemodynamic changes in the dorsolateral prefrontal cortex in a complex working memory and decision-making task.

Another study whose aim was to measure changes in mental workload at different difficulty levels by using fNIR was done by Ayaz *et* al. (2012). They used the N-back short term memory test and an air traffic controller simulation task in their study. The authors observed a monotonic decrease in reaction time and accuracy as the degree of n in n-back task was increased. fNIR results were also sensitive to this change, especially at left interior frontal gyrus. The mean oxygenation level in the arterior medial frontopolar cortex had a monotonic increase with increased task difficulty in the complex task (air traffic control task). Brain activities on different areas of prefrontal cortex according to different types of tasks (working memory and planning/decision making) were observed by fNIR. Ayaz *et* al. (2012) claim that this finding parallels with fMRI findings on the differentiation of lateral and medial prefrontal cortex in comparable tasks. To sum up, fNIR is an optical brain imaging technology that can be reliably used to measure hemodynamic changes in the prefrontal cortex and its relations with mental workload, expertise and performance in different tasks. fNIR results are in the agreement with subjective measures and earlier neuroimaging studies (Chance *et* al., 1998; Bunce *et* al., 2006; Coyle *et* al., 2007 and Izzetoglu *et* al., 2004).

When we compare fNIR with other neuroimaging techniques, we find both advantages and disadvantages of them. Firstly, EEG is relatively inexpensive, noninvasive; its results need reconstruction and can only provide a relative approximation. fMRI is safe and noninvasive neuroimaging with high spatial resolution. fMRI and fNIR are both indirect measures of neural activity limited by hemodynamic response. They are both safe and provide a level spatial resolution, but fNIR cannot substitute fMRI. When fMRI has a better spatial resolution, fNIR is limited to the outer cortex while other techniques can provide a richer picture of brain (Bunce *et* al., 2006). On the other side, fNIR tolerates movements for participants; it is more comfortable than other technologies. fNIR is safe, highly portable, user-friendly and relatively inexpensive, with rapid application times and near-zero run-time costs and these features make it a suitable tool for brain imaging in many operations compared to other brain activity measure (Izzetoglu, 2004). fNIR can be helpful to understand the cognitive and emotional state of the user during mentally demanding operations.

2.2. Cognitive Tasks

Cognitive ability is an executive function to organize a sequence of actions toward a goal (Fuster, 2008). Cognitive abilities that a brain performs are various and hard to categorize. Many studies show many different classifications. According to Deco and Rolls (2005); memory, attention and decision-making are three fundamental functions. Fuster (2008) had a longer and detailed list as following: attention, memory, working memory, planning, temporal integration, decision-making, monitoring, and inhibitory control. In our study, we followed up Deco and Rolls (2005) categorization.

Attention is a selection process by maintaining a certain amount of information from environment (Smith & Kosslyn, 2009). Working memory is a limited capacity system that involves storage, processing and manipulation of information and has an ability to encode new information into long-term memory storage (Johnson, 1992). Memory and attention are related to each other. In the industry, most employees use these functions during their daily working process. For example, air traffic controllers need to remember positions and altitudes of airplanes to prevent possible aircraft accidents. Pilots need to pay attention to every signal on display while flying. Drivers have to remember traffic signs on their road, pay attention to pedestrians and control their signals. Cognitive tasks that include memory can come up in the healthcare sector. Doctors and nurses have to remember blood pressure and other vital charts of a patient and take actions according to these vital numbers during surgery.

The Paced Auditory Serial Addition Task (PASAT; Gronwall, 1977) is a task to measure working memory and attention capacity. During the task, numbers are presented one by one every three seconds and the participants are asked at each step to sum up the last two numbers shown. This task involves both a mathematical procedure (summing) and remembering the numbers.

Operation Span, Reading Span and Counting Span, are complex span measures which are widely used for measuring working memory capacity (Daneman & Carpenter, 1980; Unsworth *et* al., 2005). In the Operation Span task, people have to both solve algebraic operations and remember unrelated words (Turner & Engle, 1989). In the Reading Span Task, sentences are presented one by one and subjects have to remember last word of each sentence and at the end of the task, they have to say all the last words in the presented order (Daneman & Carpenter, 1980). In the Counting Span task, people have to remember dots presented and count totals (Case *et* al., 1982).

N-back task is more commonly used for working memory studies (Carlson *et* al., 1998; Gevins *et* al., 1996). Participants have to decide whether a currently presented visual stimulus matches the stimulus previously presented “n” trials back during the task. N could be equal to 0, 1, 2, 3, 4, and more. As n increases, the task becomes more difficult (Parmenter *et* al., 2006). Several working memory studies, especially neuroimaging researches, use n-back task to see the role of ventrolateral and dorsolateral prefrontal cortex (Braver *et* al., 2001; Manoach *et* al., 1997; Ragland *et* al., 2002).

Decision making is a cognitive process which involves choosing one option from among a set of alternatives based on a person's criteria (Wang & Ruhe, 2007). Simon (1977) divided decision making process into three steps: intelligence (identification of all possible alternatives), design (determining results of these alternatives) and choice (evaluate these results). In the first step, a person identifies the problem, gathers information about the problem and analyses the situation. Then she/he generates possible alternatives. In the second step of decision making process, the person evaluates the results of possible options. In the third and last step she/he evaluates alternatives and selects a best alternative for her/him. In industry, it generally shows up at management departments of every sector or managerial-based sectors. For example, a director has to make decisions every day by considering his company's benefits. Doctors have to decide which medicine is better for a patient, planning engineers have to decide how many products must be produced next month, and logistic engineers have to decide where to establish their warehouses.

Some examples of the cognitive tasks used in the studies involving brain imaging techniques are as follows: Mandrick *et* al. (2013) used fNIR to see how an additional mental load affects brain activation and used an arithmetic task as a cognitive task. Cui *et* al. (2011) compared fNIR to fMRI for monitoring brain function across multiple cognitive tasks. They used four types of tasks: finger tapping, go/no-go task, judgment of line orientation task and n-back task. Bell *et* al. (2005) studied gender differences in brain activities during cognitive tasks by using fMRI. They considered a variety of cognitive task types to ensure reliable information on gender differences. These were verbal fluency, spatial attention, working memory and motor tasks. If we consider studies that used fNIR, their cognitive tasks were as follows: Bunce (2006) used target categorization, Leon-Carrion *et* al. (2010)

selected Luria’s Memory Word Task, Izzetoglu (2011) assigned target categorization as an attention task and n-back as a working memory task.

2.3. Individual Differences

2.3.1. Gender Differences

There are many studies that use brain imaging technique to investigate gender differences. Fujimoto *et* al. (2008) investigated relative metabolic changes due to age- and gender-related differences in the brain by using PET and MRI analysis. Their brain research was comprehensive; it included frontal, primary sensorimotor, parietal, occipital and temporal lobes. Since we discuss frontal cortex activities in this study, we focused on the results related to these areas. Relative metabolic values in the frontal lobe showed significant age-related differences between subjects’ at 20’s and 40’s (Their metabolic value decreases by age), but there is no significant age-related difference between subjects’ at 50’s and 70’s. Significant gender differences were not apparent in the frontal region in each age interval. Willis et al. (2002) aimed to determine the effects of age, sex and laterality on cerebral glucose metabolism (CMRglc) by using PET. The results showed that there was an inverse relationship between absolute regional CMRglc and age across widespread cortical areas including frontal cortex. Men had lower absolute metabolism than women bilaterally in the medial frontal gyrus. Both men and women showed left greater than right regional CMRglc in the medial frontal cortex while they showed right greater than left regional CMRglc in lateral frontal cortex. Marumo (2009) used an emotional activation task to see gender difference in prefrontal cortex by using fNIR. Females had significantly increased oxy-Hb change relative to males in the right ventrolateral prefrontal cortex during the latter half of the task period and this showed that the gender had an effect on individual variability of NIRS signals in response to emotional stimuli. Yang (2007) searched for gender differences in prefrontal area activation during emotional stress by using fNIR. The oxy-Hb response in the prefrontal cortex induced by emotional stress of females was significantly higher than males.

Gender differences in spatial skills is a popular topic which has been investigated by many studies. Results of these studies vary (Coluccia & Louse, 2004). There are many previous studies indicating that males show better performance than females on a diverse set of spatial tests. Galea & Kimura (1993) made an experiment about tracing a novel route through the town to find out differences between male and female subjects. Results showed that males made fewer errors and required fewer trials to learn the novel route than did females. Persson *et* al. (2013) searched for sex differences in spatial memory by using three-dimensional virtual mazes and found that men outperformed women on the maze task. Moffat *et* al. (1998) investigated sex differences in spatial ability by using computer-generated mazes. There was a significant main effect of gender. Moffat *et* al. (1998) found that males solved the mazes significantly faster than females across all five trials and males made significantly fewer errors than females. However, there are also studies that gender differences in spatial ability are totally absent. O'Laughlin & Brubaker (1998) used mental rotation and cognitive mapping tasks and found men and women performed equally well on those tasks. Taylor & Tversky (1992) also used a mental rotation test and there were no gender differences. Gender differences in visual-spatial navigation are also hot topics in neuroimaging studies. Gron *et* al. (2000) did an experiment with a maze navigation task and found out males had activations in left hippocampus whereas females had activations in right parietal and right prefrontal cortex.

There are also studies that employ the n-back task as a working memory cognitive task to investigate gender differences. Different versions of n-back task are available such as spatial, verbal and auditory. Li *et* al. (2010) found no gender differences in response time and accuracy during verbal 1-back, 2-back and 3-back tasks whereas a significant gender difference showed up in brain activations. Females had lower amplitudes and were more spatially focused than males. Koch *et* al. (2007) compared performances on verbal 0-back and 2-back tasks between males and females and observed no gender differences in both tasks. Results of Schmidt *et* al. (2009) also show that there were no significant gender differences at verbal 0, 1, 2 and 3-back tasks. Voyer *et* al. (1995) examined cognitive gender differences in the context of spatial abilities by doing a meta-analysis of many studies and they found that males outperformed females for spatial-based working memory tasks.

Speck *et* al. (2000) examined gender differences in brain activation during working memory tasks by using fMRI. 1-back and 2-back tasks were used as working memory tasks. Activation of the lateral prefrontal corticex (LPFC), the parietal corticex (PC) and caudate were observed in both sexes. Specifically, males showed right-sided dominance while females showed activation in the left hemisphere. Performance data showed that females had significantly higher accuracy and slightly slower reaction times. Both groups showed grater left inferior frontal, superior parietal and middle temporal gyrus activation. Li *et* al. (2010) worked gender differences in working memory during verbal n-back tasks by using 16 channel fNIR. They looked for changes in the concentrations of oxy-hemoglobin, deoxy-hemoglobin and total hemoglobin. While changes in oxy-hemoglobin and total hemoglobin exhibit significant gender difference, results showed that changes in deoxy-hemoglobin did not exhibit a significant gender difference, whereas females show left-lateralized activation and males showed bilateral activation for changes in oxy-hemoglobin and total hemoglobin. However, there are also studies (Schmidt, 2009; Haut & Barch 2006) which examined gender differences during n-back task and found that men and women showed similar brain activity during n-back tasks by using fMRI.

Gender differences during decision-making involves a wide concept since decision-making is a complex cognitive process including evaluation, prediction, anticipation and response. So, many different types of tasks were used as a decision making task in previous studies. Iowa Gambling Task (Bechara *et* al., 1994), Information Sampling Task (Clark *et* al., 2006), Cambridge Gambling Task (Manes *et* al., 2002) are the most known decision making tasks. There are also studies that investigate gender differences during decision making. Several studies found out that males showed better performance than females during a decision-task (Reavis & Overman, 2001; Bolla *et* al., 2004). There are also studies which indicated gender differences during a decision task and their results showed that behavior and earnings were similar for males and females (Crone et al., 2003; Lighthall *et* al., 2012). Van der Bos et al. (2012) also reached the same conclusion that there were no gender differences on deliberation time, impulsivity or completion time during a decision task.

fMRI studies on decision-making indicate that ventromedial prefrontal cortex, dorsolateral prefrontal cortex and lateral frontopolar areas show activations during decision making processes (Monchi *et* al. 2001; O'Doughtery *et* al., 2001). Bolla *et* al. (2004) investigated brain activity in the orbitofrontal cortex (OFC) and dorsolateral prefrontal cortex (DLPFC) during Iowa Gambling Task by using PET scans and found that men showed greater lateralized brain activity to the right hemisphere than females, while females showed grater brain activity in the left DLPFC. This suggests that women may have different cognitive strategies from men during a decision-making task.

Table 1 summarizes the literature review on gender differences.

Table 1: Summary of Studies on the Gender Differences

Study	N	Females	Males	Difference	Cognitive Task	Measurement	Analyses	Findings
Fujimoto *et* al. (2008)	126	62	64	Age, gender	Not stated	Brain activations by PET, MRI	t-test, regression analysis and Pearson's correlation	Significant gender difference, significant age difference between 20's and 40's, no significant age difference between 50's and 70's
Willis et al. (2002)	66	28	38	Age, gender	Not stated	Brain activations by PET	ANCOVA, Pearson's correlation	There was an inverse relationship between absolute metabolism and age and males had lower absolute metabolism in the medial frontal gyrus
Marumo (2009)	20	10	10	Gender	An emotional activitation task	Brain activations by fNIR	t-test, Pearson's correlation	Females had increased oxy-Hb change relative to males in the right ventrolateral prefrontal cortex
Yang (2007)	30	19	11	Gender	An emotional stress task	Brain activations by fNIR	ANOVA with repeated measures	Females had increased oxy-Hb change relative to males in the prefrontal cortex
Galea & Kimura (1993)	97	48	49	Gender	Tracing a novel route	Number of errors, number of trials	ANOVA with repeated measures	Males made fewer errors and required fewer trials to learn the novel route
Persson *et* al. (2013)	24	12	12	Gender	Maze tasks	Pointing errors, navigation time, pointing time, number of mazes completed, brain activations by fMRI	ANOVA with repeated measures	Males were more accurate at pointing, males were faster to complete the maze tasks, males had greater brain activitations in the right posterior and anterior hippocampus
Moffat *et* al. (1998)	74	34	40	Gender	A maze task, spatial ability tests and verbal ability tests	Performance, number of errors, completion time	ANOVA with repeated measures	Males solved the maze faster and made fewer errors during the maze, males had better performance on spatial ability tests, no sig. gender difference on verbal ability tests
O'Laughlin & Brubaker (1998)	160	82	78	Gender	A mapping task	Accuracy	2-factor ANOVA	No sig. gender difference
Taylor & Tversky (1992)	70	Not stated	Not stated	Gender	A mental rotation test	Performance	Not stated	No sig. gender difference
Gron *et* al. (2000)	24	12	12	Gender	Maze tasks	Brain activations by fMRI	t-test	Males had activations in left hippocampus, females had activations in right parietal and right prefrontal cortex

Table 1 (continued): Summary of Studies on Gender Differences

Study	N	Females	Males	Difference	Cognitive Task	Measurement	Analyses	Findings
Li *et* al. (2010)	54	28	26	Gender	Verbal 1-back, 2-back and 3-back tasks	Accuracy, response time and brain acitivitations by fNIR	ANOVA with repeated measures	Females show left-lateralized activation and males showed bilateral activation for changes in oxy-hemoglobin and total hemoglobin
Koch *et* al. (2007)	40	19	21	Gender	Verbal 0-back and 2-back tasks	Performance, brain activiations by fMRI	ANOVA with repeated measures, ANCOVA, t-test	No gender difference in performance, females showed stronger brain activations in a widespread network than males
Schmidt *et* al. (2009)	50	Not stated	Not stated	Gender	Verbal 0-back, 1-back, 2-back and 3-back tasks	Performance, brain activiations by fMRI	ANOVA with multivariate repeated measures	No sig. gender difference
Speck *et* al. (2000)	17	8	9	Gender	Verbal 1-back and 2-back tasks	Brain activations by fMRI, response time and accuracy	2-way ANOVA	Males showed right-sided dominance while females showed activation in the left hemisphere, females had significantly higher accuracy and slightly slower reaction times.
Haut & Barch (2006)	49	26	23	Gender	Episodic encoding, 2-back and yes/no recognition tasks	Brain activations by fMRI, response time and accuracy	ANOVA and t-test	No sig. gender difference
Reavis & Overman (2001)	161	95	66	Gender	Decision tasks (Iowa Card Task and California Weather Task)	Performance	ANOVA	Males had better performances
Bolla *et* al. (2004)	20	10	10	Gender	A decision task (Iowa Card Task)	Performance, brain activiations by PET	Mann Whitney U-Test	Males had better performances, males showed greater lateralized brain activity to the right hemisphere, females showed grater brain activity in the left DLPFC
Crone *et* al. (2003)	257	Not stated	Not stated	Gender	A decision task (Iowa Card Task)	Performance	ANOVA	No sig. gender difference
Lighthall *et* al. (2012)	47	23	24	Gender	A decision task (Balloon Analogue Risk Task)	Performance, brain activiations by fMRI	ANOVA, correlation tests	No sig. gender difference
van der Bos *et* al. (2012)	213	140	73	Gender	A decision task (Iowa Card Task)	Deliberation time, impulsivity and completion time	Not stated	No sig. gender difference

2.3.2. Personality Differences

Effects of personality on perceived workload are investigated by a few studies (Friedman & Rosenman, 1974; Grubb *et* al., 1994; Damos & Bloem, 1985). Friedman and Rosenman (1974) grouped personality indicators into two main types as Type A and Type B where people with Type A personality had characteristics of being impatient, competitive, aggressive, incapable of relaxation and people with Type B personality had characteristics of being relaxed, patient and easy-going. They reported Type A participants had higher scores than Type B participants on tests. Damos & Bloem (1985) found that only one between-group difference was significant under single-task conditions: Type A participants performed memory searches almost twice as quickly as Type B participants. Another important finding of this study is that Type A participants were less satisfied with their performance although they performed better in some conditions.

Sohn and Jo (2003) examined personality and its effects on mental workload to find the ideal flight crew combination. They divided their participants into four personality groups. They found a high relation between personality and NASA-TLX, which is the overall measure of mental workload.

Rose *et* al., (2002) used a broader personality trait (the Big Five personality traits) to examine relationship with both vigilance performance and perceived workload. The Big Five personality traits is a personality inventory that involves 181 items for self-reporting and generates scores for each of the five dimensions (McCrae & Costa, 1997). These dimensions are Extraversion, Agreeableness, Conscientiousness, Emotional Instability (Neuroticism) and Openness to Experience. The results of Rose *et* al., (2002) study which investigated influence of the Big Five personality traits on performance and perceived workload, showed that two of five dimensions (extraversion and conscientious) correlated with performance. With regard to perceived workload, high neuroticism associated with frustration levels.

Gencoz and Oncul (2012) adapted the Big Five personality traits to be used within Turkish culture. Their analysis on Turkish people presented us a new personality characteristic which was called Negative Valence dimension and added to the five dimensions in the original the Big Five personality traits. Their study created a Basic Personality Traits Inventory (BPTI) with six personality dimensions (extraversion, neuroticism, agreeableness, conscientiousness, openness to experience and negative valence) identified for the Turkish culture.

Since one of the aims of this thesis study is to investigate whether personality has an effect on mental workload during different cognitive tasks, using a generally accepted personality test is important. Since our experiments were planned to be done in Turkey, Gencoz and Oncul (2012)'s study was very helpful and their BPTI with six dimensions was used in this study.

There are few studies used personality traits with brain imaging techniques. DeYoung (2010) used Big Five personality traits to investigate the association of each trait with different brain areas. As a result, four of five traits (extraversion, neuroticism, agreeableness and conscientiousness) supported their hypothesis. Extraversion has a significant positive association with the activations in medial orbitofrontal cortex. Neuroticism is negatively correlated with the activations in right dorsomedial PFC and in portions of the left medial temporal lobe, positively correlated with the activations mid-cingulate cortex. Agreeableness is associated positively with the activations in the retrosplenial region of posterior cingulate cortex and fusiform gyrus, associated negatively with the activations in superior temporal sulcus. Lastly, conscientiousness has a significant positive association with the activations in lateral PFC and a significant negative association with the activations in posterior fusiform gyrus.

There are many types of personality traits and different kinds of cognitive tasks and hence the relationship between personality and cognitive ability is a quite wide research area. However, results of relationship between cognition tasks and personality are not always consistent. Extraversion, conscientiousness and agreeableness change according to type of cognitive task. Extraversion was positively correlated with speed and memory, but it was negatively correlated with verbal ability (Chamorro-Premuzic *et* al., 2006). According to Costa *et* al. (1976), high anxiety was associated with lower cognitive performance, while high openness to experience and low extraversion were associated with higher cognitive scores. Graham & Lachman (2012) found that personality stability was associated with better cognitive performance. Conscientiousness had both positive and negative relationships with cognitive tasks (Furnham & Chamorro-Premuzic, 2006; Moutafi *et* al., 2006). Results of Graham and Lachman (2014) showed that openness was positively correlated with verbal fluency, extraversion was negatively correlated with reasoning, neuroticism was negatively correlated with reasoning. Judge *et* al. (1999) claimed that high extraversion, high conscientiousness, low agreeableness and low neuroticism were associated with extrinsic career success, when high conscientiousness was associated with intrinsic career success.

Specifically, Davies (1965) used two Heron subscales (which measure emotional stability and sociability) as a personality trait during his experiment and found that there was no significant correlation between maze test scores and personality. The Big Five personality traits has recently been used in many studies that focus on the effects of personality type on decision-making tasks that involves delay discounting, reward sensitivity, gambling, and risk-taking since it was first used by Lauriola and Levin (2001). According to Lauriola and Levin (2001)'s results, people with low neuroticism and people with high openness to experience took more risk in a decision making task, while the correlation coefficients were not significant between risk-taking and other personality traits (extraversion, conscientiousness, agreeableness). Hooper *et* al. (2008) found that there was a negative relationship between neuroticism and performance in a decision-making task. Byrne (2015) used the Big Five personality traits to examine the relationship between a decision making task and personality and he found that agreeableness, conscientiousness and neuroticism negatively correlated with performance in the decision making task.

Table 2 summarizes the literature review on personality differences.

Table 2: Summary of Studies on Personality Differences

Study	N	Females	Males	Personality Test	Cognitive Task	Measurement	Analyses	Findings
Friedman &Rosenmann (1974)	Not stated	Not stated	Not stated	Type A and Type B	Cognitive tests	Performance	Not stated	Type A participants had higher scores
Damos & Bloem (1985)	16	16	0	Type A and Type B	Single and dual-tasks	Reaction time and accuracy	Not stated	Type A participants performed faster
Sohn and Jo (2003)	61	Not stated	Not stated	Myers Briggs Type Indicator	4 – 6 flight	Heart rate, altitude deviation, and NASA-TLX	Two-way MANOVA, one-way MANOVA	A high relation between personality and NASA-TLX
Rose *et* al., (2002)	96	48	48	The Big Five personality traits	A vigilance task	Performance, false alarm, reaction time and NASA-TLX	One-way ANOVA, mixed ANOVA, partial correlation	Extraversion and conscientious correlated with performance, neuroticism correlated with workload
DeYoung (2010)	116	58	58	The Big Five personality traits	Not stated	Brain activations by fMRI	Correlation and regression-based analyses	Brain activations correlated with extraversion, neuroticism, agreeableness and conscientiousness traits
Chamorro-Premuzic *et* al. (2006)	118	87	31	Eysenckian personality	Verbal and numerical cognitive tasks	Performance	Correlation and regression-based analyses	Verbal ability was associated with introversion, dissimulation and neuroticism. Numerical ability was associated with caution
Costa *et* al. (1976)	969	0	969	Anxiety, Extraversion and Openness to Experience	Information Processing Ability (IPA), Manual Dexterity (MD), and Pattern Analysis Capability (PAC)	Performance	Correlation and regression-based analyses	High anxiety was associated with lower performance in all tasks, high openness to experience was associated with higher IPA and PAC scores, low extraversion was associated with higher PAC
Graham & Lachman (2012)	4974	Not stated	Not stated	Personality Stability	Cognitive tasks	Performance and reaction time	Multiple regression and ANCOVA	Personality stability was associated with better cognitive performance
Furnham & Chamorro-Premuzic (2006)	93	70	23	The Big Five personality traits	Examinations, a cognitive ability and beliefs about intelligence test	Examination grades, cognitive ability performance and beliefs about intelligence performance	Partial correlation, ANOVA	Conscientiousness, extraversion significantly correlated with examination grades. There was a correlation between conscientiousness and beliefs about intelligence performance.

Table 2 (continued): Summary of Studies on Personality Differences

Study	N	Females	Males	Personality Test	Cognitive Task	Measurement	Analyses	Findings
Moutafi *et* al. (2006)	2658	Not stated	Not stated	The Big Five personality traits	Cognitive tasks	Fluid intelligence	Correlational analyses	Openness was positively correlated with fluid intelligence, while conscientiousness was negatively correlated with it.
Graham and Lachman (2014)	154	73	81	The Big Five personality traits	A category fluency task, generating number, counting and Stop and Go task.	processing speed, reaction time,verbal fluency, reasoning, memory	Hierarchical multiple regression	openness positively correlated with verbal fluency, extraversion negatively correlated with reasoning, neuroticism negatively correlated with reasoning.
Judge *et* al., 1999	354	Not stated	Not stated	The Big Five personality traits	Jon satisfaction, status, income,	intrinsic success and extrinsic success	Correlation and regression-based analyses	High extraversion, high conscientiousness, low agreeableness and low neuroticism were associated with extrinsic career success, when high conscientiousness was associated with intrinsic career success
Davies (1965)	540	240	300	the Heron Inventory	The Perceptual Maze Test	Performance	Product-moment correlation	No sig. personality difference
Lauriola and Levin (2001)	Not stated	Not stated	Not stated	The Big Five personality traits	the Ambiguity-Probability Tradeoff Task	Risk-taking performance	Correlation and regression-based analyses	Low neuroticism and high openness to experience were associated with greater tisk taking
Hooper *et* al. (2008)	Not stated	Not stated	Not stated	Eysenckian personality	A decision task (Iowa Card Task)	Performance	Multiple regression analyses	There were a relationship between neuroticism and performance and a relationship between extraversion and performance
Byrne (2015)	127	76	51	The Big Five personality traits	A dynamic decision making task	Completion time and performance	t-test, correlational analyses and two-step hierarchical multiple regression	Agreeableness, conscientiousness and neuroticism negatively correlated with performance in the decision making task
Gray (2001)	152	76	76	Extraversion and neuroticism	Spatial or verbal 2-back task	Reaction time and accuracy	mixed ANOVA	Extraversion and neuroticism did not have any correlations with reaction time and accuracy
Canli *et* al. (2001)	Not stated	Not stated	Not stated	The Big Five personality traits	Emotional experience	Brain activations by FMRI	Correlation and regression-based analyses	Brain activations positively correlated with extraversion and negatively correlated with neuroticism.

CHAPTER 3

THE METHODOLOGY

In this chapter, information about participants and apparatus are given. Cognitive tasks used in experiments and procedures for experiments are discussed in detail. Before experiments, all participants signed the Informed Consent Form approved by the Human Subject Ethics Committee at the Middle East Technical University as seen in Appendix A.

3.1. Participants

Participation to the study was voluntary. 37 right-handed participants (23 female, 14 male) were recruited for this study. Their mean age was 25.51. They were undergraduate or graduate engineering students at the Middle East Technical University or Hacettepe University. Demographical data of participants is shown in Appendix B.

3.2. Apparatus

This section includes the description of materials, software and systems used during the experiments. In this study a questionnaire about user information, the Informed Consent Form, the Basic Personality Traits Inventory, fNIR Device, an n-back task, a maze task, an information sampling task were used. In the following sections these apparatus will be explained in detail.

3.2.1. Functional Near-Infrared (fNIR) Spectroscopy

fNIR technology is sensitive to mental workload which is provided by measuring the change in the rate of oxygenated and deoxygenated hemoglobin in the prefrontal cortex. Continuous wave NIR spectroscopy system supplied by fNIR Devices LLC (Ptomac, MD: www.fnirdevices.com) was used during all experiments. The system included a control box, two sensor pads (each piece contains 2 channels) and a COBI Control Device Software (Figure 1). These two sensor pads were designed to monitor dorsal and inferior frontal cortical areas of the brain (Ayaz *et* al., 2012), so they were set up on right and left area of the forehead respectively (Figure 2). Data acquisition, collection and presentation were provided by Cognitive Optical Brain Imaging Studio (COBI) Control Device and fNIRSoft Software (Ayaz *et* al., 2012). COBI is a software where users are able to acquire, process and visualize fNIR signals, whereas fNIRSoft offers tools to filter and analyze raw fNIR signals.

Figure 1: fNIR System Design

Figure 2 - fNIR Sensors Placed on the Participant's Forehead

On each sensor pad, there is a light source which emits light changes between 700-900 nm wavelengths. The light in this frequency band is primarily absorbed by hemoglobin molecules, so changes in the concentration of oxy-hemoglobin and deoxy-hemoglobin can be observed by the sensors through the channels (Ozcan, 2012). After measuring hemodynamic changes, analyses were performed by using fNIRSoft software. The increase in oxy-hemoglobin concentration with respect to deoxy-hemoglobin concentration indicates brain activation and functional challenges (Izzetoglu, 2004). Therefore, fNIR will be helpful for measuring mental workload.

3.2.2. Information Sampling Task

The Information Sampling Task (IST) was used as a decision making cognitive task. The IST was presented for the first time by Clark *et* al. (2006).

The IST is a measure to gather information about people's tendency to make a decision. The task involves a five-by-five matrix of closed boxes and opened forms of these boxes are one of two colours. The aim of the task is finding which colour is dominant to the other one. Subjects can open these boxes by clicking on them and they are free to open as many as they want to decide which colour is the majority in the 5x5 matrix. When subjects decide on the dominant colour in the matrix at some point, they press on one of the two colours in the screen. The process can be seen in Figure 3.

There are two levels of the task with 10 trials in each of them: Fixed win (FW) and decreasing win (DW). In the FW, subjects start with 0 points to the task, they can open as many boxes as they want and number of opened boxes does not affect their points. After their decision, if it is correct they win 100 points. Otherwise, they lose 100 points. In the DW level, subjects start with potential 250 points, and every box opened by subjects decreases the amount of the possible award by 10 points. At the decision point, if their decision is correct, they win the leftover of 250 points; if it is not, they lose 100 points. The equations used to calculate for IST-1 and IST-2 performances are as follows, respectively:

$$A_{1=} \sum_{i=1}^{10} 100\, D_i \ , \qquad \text{Eq.1}$$

$$A_{2=} \sum_{i=1}^{10} ((250 - n_i)\, K_i - 100(1 - \ K_i)), \qquad \text{Eq.2}$$

where D_i is equal to 1 if the decision of the participant is correct at trial-i during IST-1 and D_i is equal to -1 if it is not. For IST-2, K_i is used for accuracy. K_i is equal to 1 if the decision of the participant is correct at trial-i during IST-2 and K_i is equal to 0 if it is not.

The script for the IST was supplied by Millisecond Software LLC (www.millisecond.com) and it was run on a computer during the experiment.

Clark *et* al. (2006) used trial numbers (numbers of boxes opened), errors, latency of box opening and total points for his analyses and results. In this study, trial numbers, completion time of the task and performance were used for analyses.

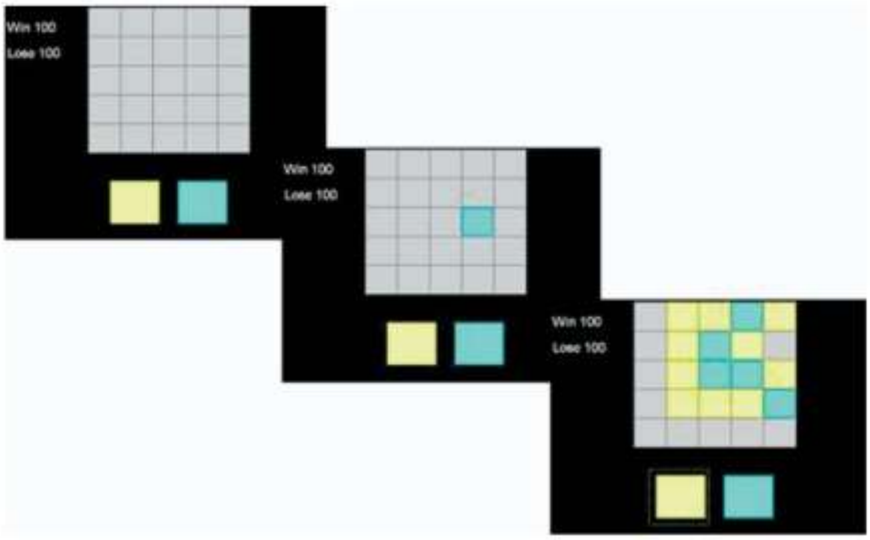

Figure 3: Screen Display for the IST (Clark *et* al., 2006)

3.2.3. N-Back Task

N-back task is a standard attention and working memory task (Watter, 2001). 3x3 matrix was shown with an object which was occupied in any position in matrix. The stimuli duration was 500 milliseconds and there were 2500 milliseconds until the next stimuli. At each stimulus, the object could be at any place in the matrix. In this study, four conditions were used from one to four to create different difficulty levels. As each stimulus was presented, the participant had to compare the position of the current stimulus with the stimulus that occurred n items before. The participant was asked to keep track of the position of the object and in the 1-back condition, if the position of the object was identical to the one presented previously, the participant had to press ‘A’ button. In the 2-back, 3-back and 4-back condition, if the object’s position was identical to the one presented two, three and four trials back, respectively, the participant had to press ‘A’ button. Each condition had 20 trials. A screen display of 2-back condition can be seen in Figure 4.

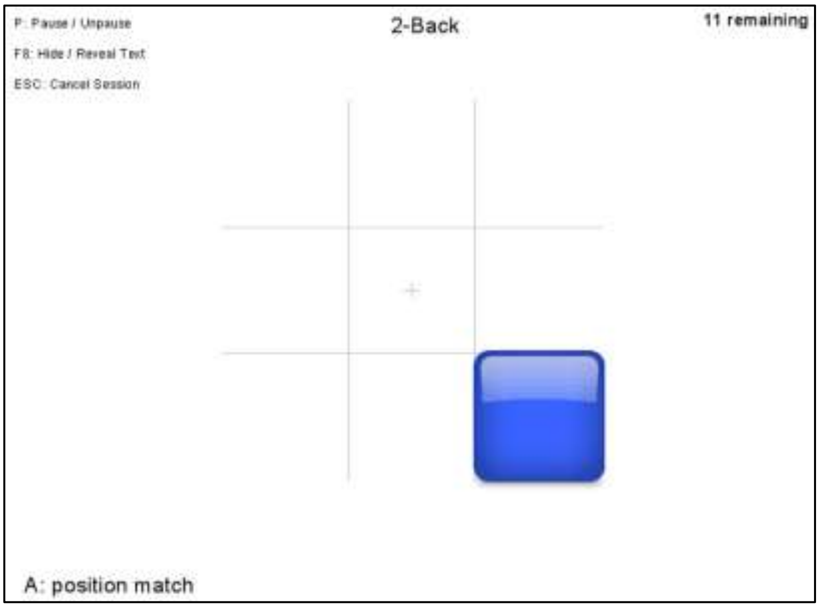

Figure 4: Screen Display for N-back Task

Brain Workshop 4.8 was used for n-back task which was supplied from Brain Workshop web page (http://brainworkshop.sourceforge.net/). The program gave accuracy results of each n-back. N-back accuracy was calculated by the following equation:

$$A_3 = \frac{M}{P + U} \quad , \qquad \text{Eq.3}$$

where M is the number of pressing the button when position is matched, U is the number of pressing the button when position is unmatched, and P is the number of total same positions of objects.

The completion time was the same for every participant, because stimuli had same trials in each experiment. Therefore, only accuracy of each task was used as a performance measure.

3.2.4. Maze Task

Maze is a task that demands visual-spatial skills (Ayaz *et* al., 2008). Maze contains both working memory function to remember where you are and decision-making task to decide turning to right or left. Spatial ability is a very common study area on gender differences. Therefore, maze is chosen as another cognitive task to be used in our study.

Maze Suite application was used as a maze task. A 3-D maze environment was created by using Maze Suite which was first described and presented by Ayaz *et* al. (2008), copyrighted by Drexel University and obtained from Maze Suite webpage (mazesuite.com/downloads). It is an application to create 3D virtual environments. It is used for researches based on navigational and spatial cognitive neuroscience experiments. Maze Suite is composed of three applications: MazeMaker, MazeWalker and MazeAnalyzer. MazeMaker is the editor application to create and edit experiments. MazeWalker is the application that renders mazes created by MazeMaker. Lastly, MazeAnalyzer is the application for analysis.

Having full control of designing, running and analyzing in one application is a benefit for our spatial based experiment. Another advantage of Maze Suite is its ability to send signals to fNIR through a computer serial port which is an advantage for synchronizing fNIR measures with behavioral performance data.

3.2.4.1. Maze Maker

Maze Maker is the first phase of Maze Suite application. 3D maze environments are designed in this step.

The path, start and finish points, any different objects on the path were created in there. The maze should not have been too easy since the focus of this study is mental workload. But also, we did not want to design a very complicated maze that people would struggle to finish very hardly and maybe give up by considering all other cognitive tasks. The top view of the maze design used in this study is shown in Figure 5.

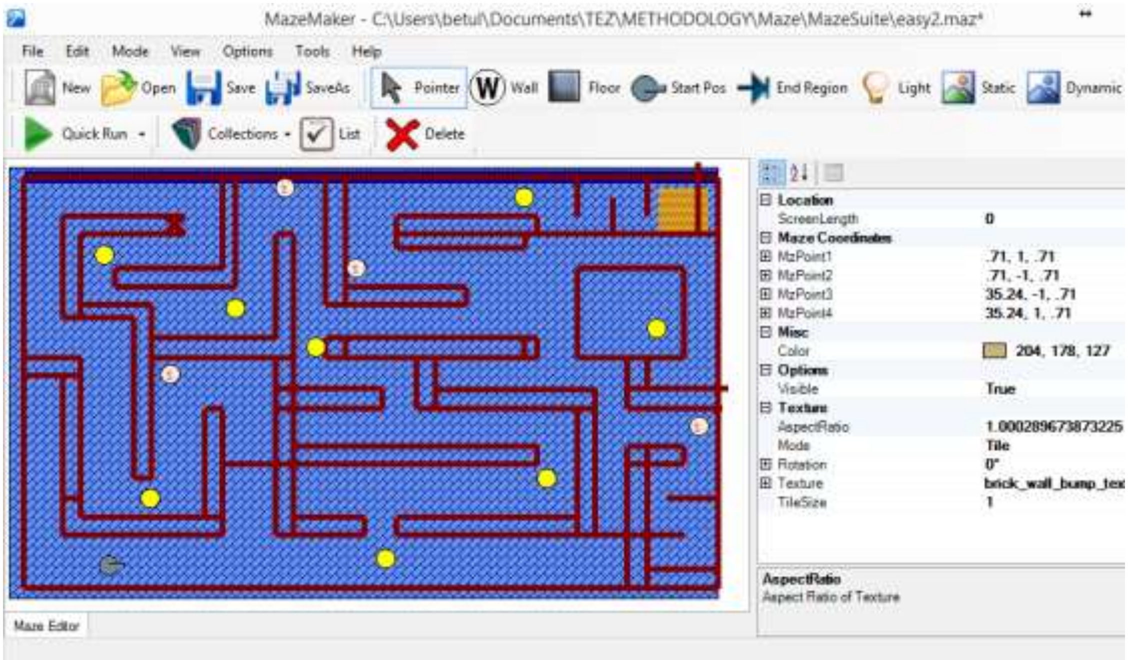

Figure 5: Design of the Maze in Maze Walker

3.2.4.2. Maze Walker

Maze Walker is the enabler of Maze Suite (Ayaz *et* al, 2008). The maze that will be used is chosen and started with Maze Walker. An initialization dialog box of this application provides for arranging initial settings of the maze such as video (resolution, colour depth, full screen), control (use mouse/joystick' use pre-recorded log file), visual (show crosshair/timer/bar, enable lights/shaders, skip warnings), maze (selection of maze) and logging (automatically save a log file, enter the name of user) (Figure 6). After all settings are done, users are able to start the maze by clicking on "Start" on this application.

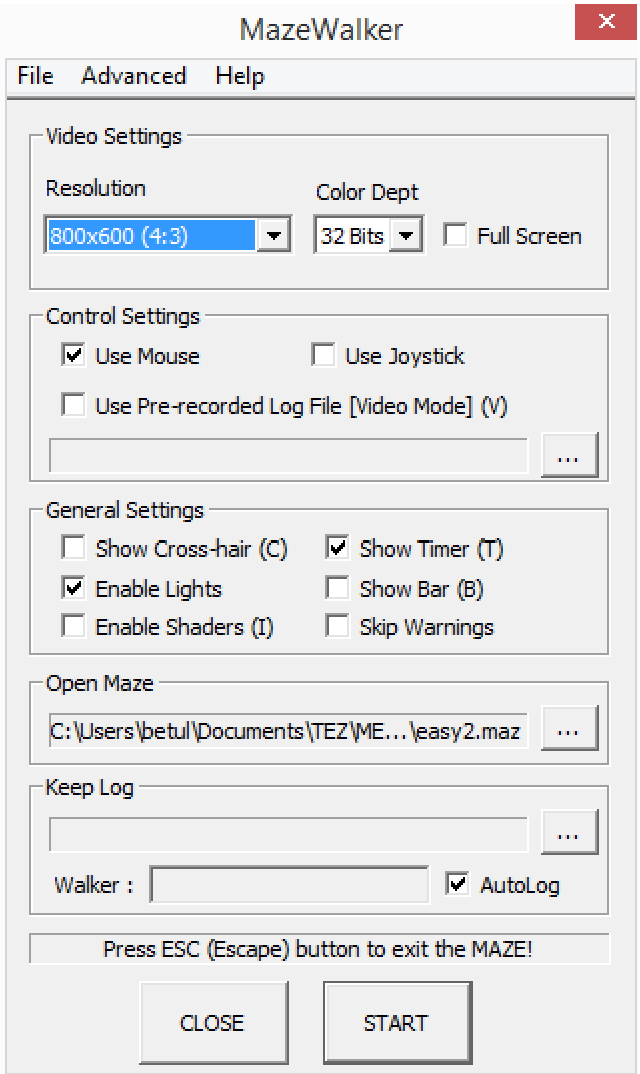

Figure 6: Initial Setting Screen of Maze Walker

In this study, resolution was set up 800x600, colour depth was chosen highest resolution (32 Bits) and full screen option was checked to provide full concentration. Keyboard use during maze was requested from users and mouse use was forbidden to ensure equal condition. The maze that was opened was the one whose design is shown in Figure 6. Auto log was checked to ensure all log files would be recorded because they would be used in analysis phase.

3.2.4.3. Maze Analyzer

The analysis phase of Maze Suite is Maze Analyzer. Log files that are recorded during maze task can be opened and user behavior can be investigated. Completion time of the maze, length of the path that is toured until finish line and visual of the path that is travelled in the maze map can be given by Maze Analyzer after the task is completed as seen in Figure 7. In this study, completion time and path length were used as performance measures.

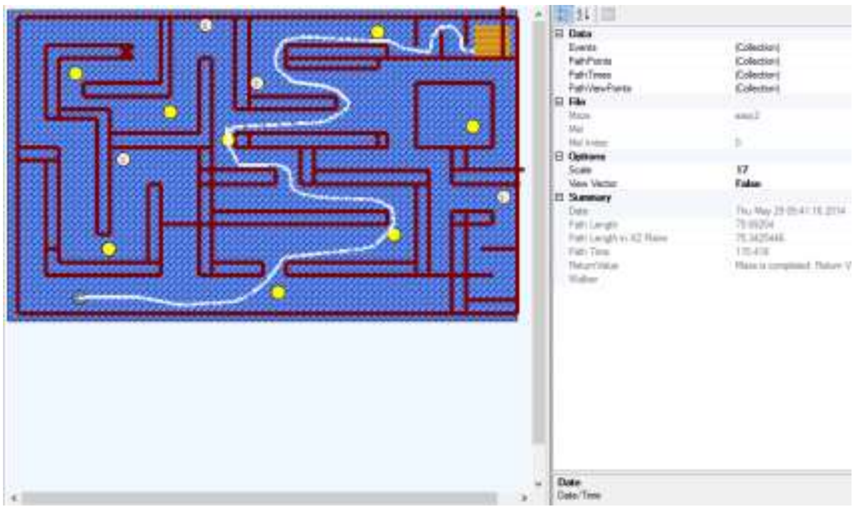

Figure 7: Maze Analyzer

3.2.5. Basic Personality Traits Inventory

After three cognitive tasks were completed, participants were asked to complete the Basic Personality Traits Inventory (BPTI). BPTI with six dimensions which is validated for use with the Turkish population (Gencoz and Oncul, 2012) was used as a personality trait and 45 items were ranked by participants. The questionnaire and English version of the questionnaire can be seen in Appendix C and D, respectively. This was the last step of the procedure.

These 45 items gave results under six dimensions which were extraversion, conscientiousness, agreeableness, neuroticism, openness to experience and negative valence. At each item, participant was asked to rate his/her familiarity with the adjective from 1 to 5. There were 8 items for extraversion, 8 for conscientiousness, 8 for agreeableness, 9 for neuroticism, 6 for openness to experience and 6 for negative valence dimensions respectively. Arithmetic means for each dimension was calculated by using items belonging themselves. Some of items are reversing entry, so their point was taken as six minus their actual point. Then, arithmetic mean calculation was done. Which items belong to which dimension is shown in Appendix E. The results of the inventory gave an idea about how the person shows similarity with the dimension in a scale of 1 to 5. In this study, performance and fNIR measures were analyzed for each dimension.

3.3. Experimental Procedure

The experiments were conducted at the METU Informatics Institute. Experimental setup included an fNIR system and a computer to run all the tasks.

The aim of this study, its procedure and tasks were explained to all of the subjects orally. User information questionnaire was applied before the experiment to obtain demographics data from the participants, which can be seen in Appendix B. Then, the participants read and signed the Informed Consent Form (Appendix F). Later, sensor pads of fNIR system were placed on the participants' forehead, signal quality check was performed and baseline measures were obtained.

SAN Suite software was used for automated execution of the experimental protocol. The break between the tasks were 15-20 seconds. After their comfort was ensured, the other phase of the experiment was started. The experiment had three phases (n-back, IST, maze) and they were chosen randomly to prevent sorting effects.

During all three tasks, fNIR pads were on the participants' forehead and were not removed. After completing each task, participants were asked if the pads were giving them any discomfort, and if so the fNIR pads were removed for a short duration to give them some relief. Participants were informed that they can quit the experiment at any time they wish due to any discomfort they may experience. However, nobody felt that way and there were not any incomplete experiment. Following the completion of the cognitive tasks, the Turkish culture based basic personality traits inventory based on 45 items was completed by participants.

N-back task took 5-6 minutes in length with a break between levels. The total time of the IST changed from person to person, but it was typically between 7 and 15 minutes. Maze task lasted between 1 and 6 minutes. Overall, completion of the experiment took 45 minutes on average, including the initial introduction, all tasks and the questionnaires.

CHAPTER 4

RESULTS AND DISCUSSIONS

In this chapter, results of analyses are given and discussed.

4.1. Data Analysis

Data of 37 participants (23 females and 14 males) were included for the analysis of performance measures. However, due to excessive noise and motion artifacts some of the subjects had to be excluded from the fNIR analysis. In particular, data from 19 participants (8 females and 11 males) could be used for the fNIR analysis.

There are 2 detectors on each fNIR sensor and each of them provides data on changes in the hemoglobin and oxy-hemoglobin levels. The raw fNIR Data (4 channels x 2 wavelengths) were filtered by FIR (Finite Impulse Response Digital Filter) to decrease high frequency noise, respiration and cardiac cycle effects (Ayaz *et* al., 2012). An approach SMAR (Sliding- Window Motion Artifact Rejection) (Ayaz *et* al., 2010) was performed for motion artifact detection and rejection from the refined fNIR measures. fNIRSoft was used to filter the data and calculate the blood oxygenation levels. fNIR measures data of 19 participants out of 37 participants were available to be used after filtering data by FIR and SMAR to eliminate high frequency noise, respiration, cardiac cycle effects, saturation and head motion artifacts. Since activations in voxel-1 and voxel-4 were enough for cognitive tasks, we decided to use them.

This study used maze completion time, maze path length, IST-1 performance, IST-1 trial number, IST-1 completion time, IST-2 performance, IST-2 trial number, IST-2 completion time, 1-back performance, 2-back performance, 3-back performance, and 4-back performance as performance measures. The study also used maze oxygenation levels in voxel-1 (left DLPFC) and voxel-4 (right DLPFC); IST-1 and IST-2 oxygenation levels in voxel-1 (left DLPFC) and voxel-4 (right DLPFC); 1-back, 2-back, 3-back, and 4-back oxygenation levels in voxel-1 (left DLPFC) and voxel-4 (right DLPFC) as fNIR measures.

SPSS 20 for Windows (SPSS Inc., Chicago,IL, USA) was used for data analysis. First, boxplots and histograms of variables were drawn for descriptive analysis of collected data. There were outliers on maze completion time, maze path length, IST-1 completion time, IST-2 completion time, IST-2 trial number. First, reasons of these outliers were investigated. However, all data were entered correctly and participants did not have any unexpected situations. There was not any problem in collecting data, so eliminating outliers was out of option and transformations of these outliers were considered. Since maze completion time, maze path length, IST-1 and IST-2 completion time show positive skewness; square root transformations were applied (Field, 2009). However, data transformations were not able to eliminate the outliers. Only one solution was left for outliers: changing the score. The mean plus two times the standard deviation was applied. There was also an outlier in trial number on IST-2 data and it did not suit any form of transformation rule. Since there was not any special case and deleting the value seemed losing a data without a reason, changing the score was decided. The mean plus two times the standard deviation was also applied for that data set (Field, 2009). After these changes, new boxplots of maze and IST performance measures were drawn. All boxplots of performance and fNIR measures are shown in the appendices. Boxplots for maze completion time, maze path length, IST performances, IST completion times, IST trial numbers, n-back performances can be seen in Appendix G, H, I, J, K and L, respectively. fNIR measures during Maze task, ISTs and n-back tasks are given in Appendix M, N and O, respectively.

All data were examined for normality by using the Shapiro-Wilk test. While all fNIR data are normally distributed (Table 3), some performance variables violated the assumptions of normality (Table 4). In Table 3, Shapiro-Wilk tests for all variables are non-significant ($p > .05$) which shows that these samples are not significantly different from a normal distribution. However, Table 4 shows that the distributions of maze path length, maze time, IST-1 performance, IST-2 completion time, 1-back performance and 2-back performance are significantly different from a normal distribution ($p < .05$). When the assumptions of normality are not satisfied, parametric tests cannot be used. Therefore, nonparametric tests were applied for these variables. The flowchart of data analysis steps is given in Figure 8.

Table 3: Normality Tests for fNIR Variables

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
V1_maze	0.133	19	.200*	0.939	19	0.254
V4_maze	0.131	19	.200*	0.977	19	0.897
V1_ist1	0.101	19	.200*	0.973	19	0.829
V4_ist1	0.136	19	.200*	0.966	19	0.694
V1_ist2	0.151	19	.200*	0.957	19	0.517
V4_ist2	0.117	19	.200*	0.982	19	0.966
V1_1back	0.123	19	.200*	0.976	19	0.891
V4_1back	0.15	19	.200*	0.964	19	0.644
V1_2back	0.124	19	.200*	0.973	19	0.834
V4_2back	0.113	19	.200*	0.982	19	0.962
V1_3back	0.113	19	.200*	0.981	19	0.952
V4_3back	0.09	19	.200*	0.966	19	0.686
V1_4back	0.09	19	.200*	0.98	19	0.944
V4_4back	0.138	19	.200*	0.97	19	0.773

*. This is a lower bound of the true significance.

Table 4: Normality Tests for Performance Variables

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
MazePath	0.147	37	0.043	0.94	37	0.047
MazeTime	0.114	37	.200*	0.938	37	0.041
IST1Per	0.299	37	0	0.791	37	0
IST1Time	0.109	37	.200*	0.95	37	0.094
IST1TrialNo	0.09	37	.200*	0.953	37	0.118
IST2Per	0.095	37	.200*	0.984	37	0.869
IST2Time	0.195	37	0.001	0.898	37	0.003
IST2TrialNo	0.1	37	.200*	0.957	37	0.157
back1	0.43	37	0	0.638	37	0
back2	0.155	37	0.024	0.898	37	0.003
back3	0.083	37	.200*	0.978	37	0.66
back4	0.082	37	.200*	0.968	37	0.367

*. This is a lower bound of the true significance.

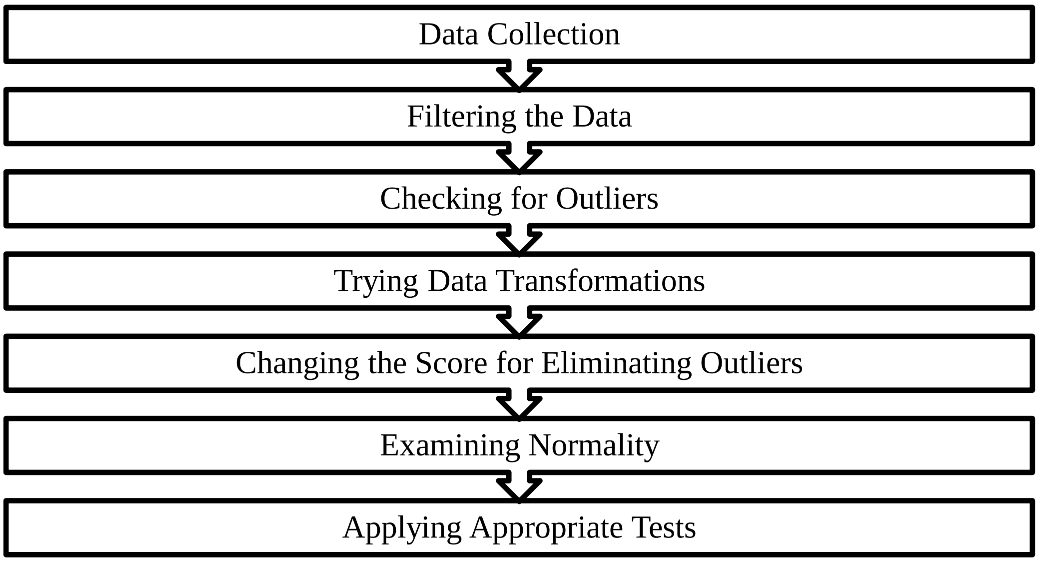

Figure 8: Flowchart of Data Analysis

4.2. Gender Difference in Performance Measures

Mann Whitney U-Tests were conducted for 37 participants (23 females and 14 males) to see whether there was gender difference in performance during maze task for each maze performance measure: maze completion time and maze completion length. Results showed that the mean maze completion time did not significantly differ between males (M=188.79, SD=91.855) and females (M=255.45, SD=116.510) where U=104.5, z=-1.770, p>.05. Figure 9 shows average values for males and females in maze completion time. Summary of the test can be found in Table 5.

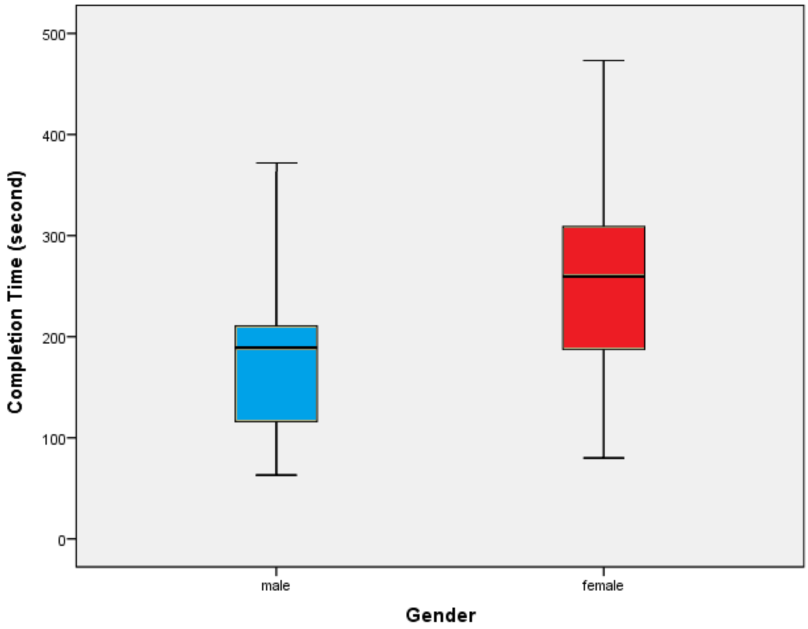

Figure 9: Boxplot of Maze Completion Time according to Gender

For another maze performance measure, results showed that maze path length did not significantly differ between males (M=195.75, SD=125.960) and females (M=193.20, SD=125.110) where U=156, z=-0.157, p>.05. Figure 10 shows average female and male values in maze path length. Summary of the test can be found in Table 5.

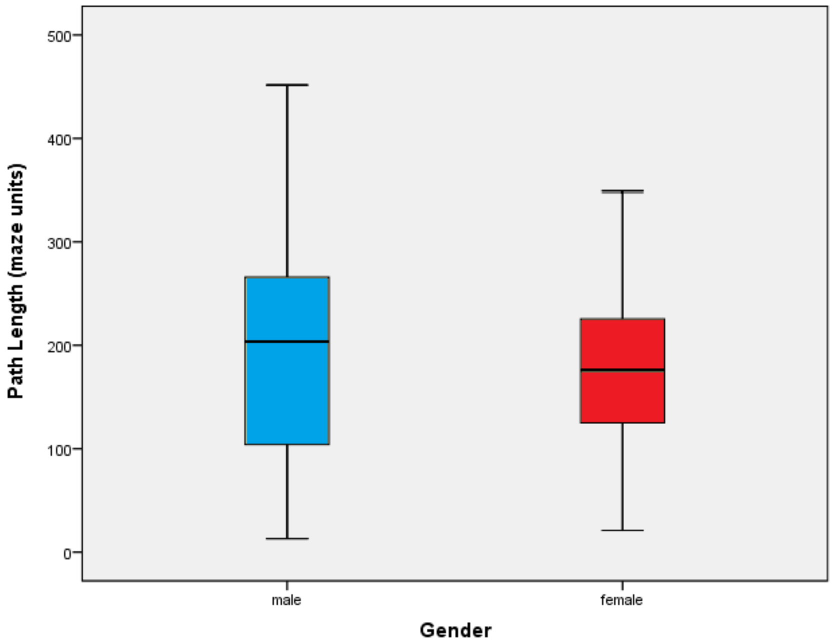

Figure 10: Boxplot of Maze Path according to Gender

Table 5: Mann Whitney U-Test Summary for Maze Performance Measures

	Maze Completion Time	Maze Path Length
Mann-Whitney U	104.500	156.000
Wilcoxon W	209.500	432.000
Z	-1.770	-.157
Asymp. Sig. (2-tailed)	.077	.876
Exact Sig. [2*(1-tailed Sig.)]	0.077	0.889

As a result it can be concluded that there is not a significant gender difference in maze performance measures. There are many studies about gender differences on spatial ability and the results of these studies range from males outperforming females (Galea & Kimura, 1993; Moffat, 1998; Waller *et* al., 2001) to no gender differences (O'Laughlin & Brubaker, 1998; Taylor & Tversky, 1992; Sadalla & Montello, 1989; Montello & Pick, 1993).

When we considered only maze-based tasks and excluded other kinds of spatial ability tests, the studies (Galea & Kimura, 1993; Persson, 2013; Moffat, 1998 indicate males are significantly better than females on maze tasks. However, we found that performances of females and males are not significantly different from each other in our study. There may be two reasons for this result. First reason of this mismatch may be based on the demographic properties of the participants. The participants of previous studies were students who enrolled in a psychology course, whereas the participants of our study were undergraduate or graduate students in an engineering degree program. People in our study were more likely to have spatial perception skills since spatial-visual skills are essential for success in engineering. Some of the engineering courses may increase students' ability on spatial and visual tasks, especially design based courses; therefore engineering students may have higher spatial ability. That fact may be the reason for why there was no a gender difference in the maze task.

Second, the maze task in our study was an easier and shorter than the maze tasks used in other studies (Moffat, 1998; Waller, 2001). In our study there were other cognitive tasks to be completed by the participants, so we opted for shorter maze designs. In our study there was one trial that lasted between 2 to 7 minutes. Moffat (1998) used 4 practice trials and 5 trials on each of two experimental mazes. Similarly, a two-hour long experiment was employed in Waller (2001)'s study. Therefore, using longer maze tasks or increasing the number of trials may underlie the gender difference reported in those studies. A reorganized maze according to these conditions can be applied for future studies.

For the information sampling task, IST-1 performance and IST-2 completion time did not satisfy the assumptions of normality as shown before, therefore a Mann Whitney U-Test (a nonparametric test) was conducted for 37 participants (23 females and 14 males) to examine whether there was gender difference in completion time / performance during IST-1 Task / IST-2 Task. Results showed that the mean IST-1 completion time did not significantly differ between males (M=184.44, SD=64.92) and females (M=146.19, SD=55.89) where U=109.5, z=-1.613, p>.05. For the second phase of the IST, the mean IST-2 completion time did not significantly differ between males (M=184.55, SD=62.91) and females (M=146.38, SD=70.88) where U=102.5, z=-1.832, p>.05. Figure 11 shows boxplots for males and females in IST-1 and IST-2 completion time. Summary of the test can be found in Table 6.

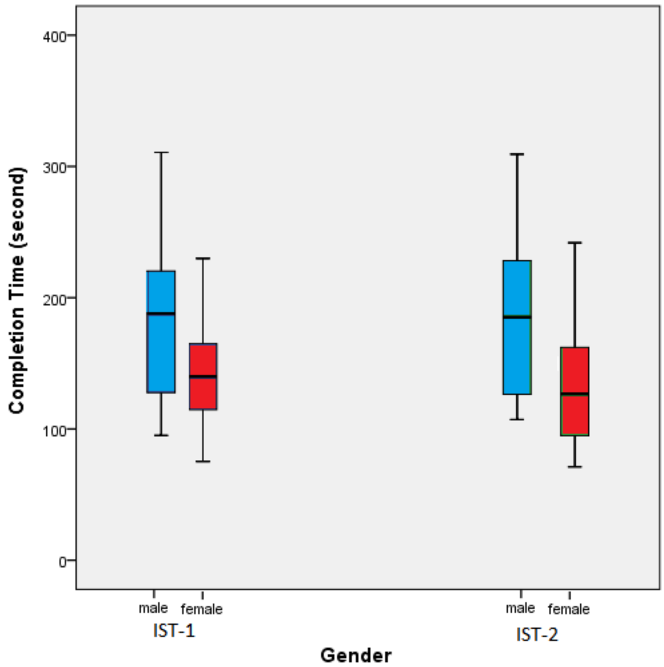

Figure 11: Boxplot of IST Completion Times according to Gender

Briefly, results show that there is no gender difference in completion time during both information sampling tasks. This finding is compatible with the literature (Crone et al., 2003; Lighthall *et* al., 2012). Completion time of a decision-making task did not change according to gender in previous studies.

A Mann Whitney U-Test was applied for 37 participants (23 females and 14 males) to investigate the gender difference in performance of IST-1 and IST-2 tasks. Result showed that the mean IST-1 performance did not significantly differ between males (M=900.00 SD=188.11) and females (M=765.22, SD=205.84) where U=100, z=-2.055, p>.05. Similarly, for the second phase of the IST, the mean IST-2 performance did not significantly differ between males (M=376.00, SD=114.37) and females (M=414.09, SD=176.56) where U=136.5, z=-0.767, p>.05. Figure 12 shows average values for males and females in IST-1 and IST-2 performance. Summary of the test can be found in Table 6.

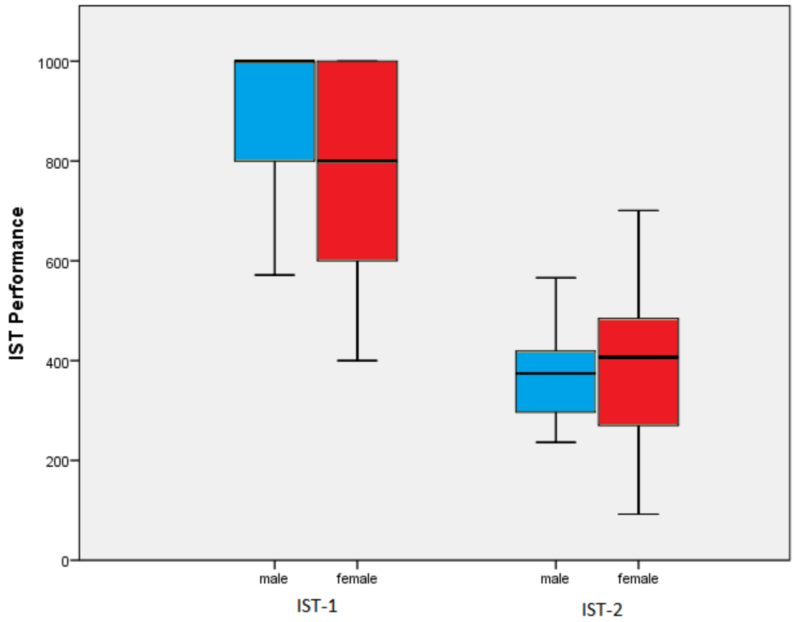

Figure 12: Boxplot of IST Performances according to Gender

Previous studies involve two kinds of results: males outperformed females in decision-making tasks (Reavis & Overman, 2001; Bolla *et* al., 2004) and no gender differences during a decision-making task (Crone *et* al., 2003; Lighthall *et* al., 2012). The results of this study show that there is no gender difference in completion time during both information sampling tasks. The participants of this study were engineering students at well-known universities which means they performed similarly at the college admissions test and they experienced a similar training in engineering. Therefore, their cognitive abilities are possibly very close to each other which may be the reason od observing no gender difference in this sample.

Table 6: Mann-Whitney U-Test Summary for IST Performance Measures

	IST1- Time	IST2-Time	IST1 - Performance	IST2 - Performance
Mann-Whitney U	109.500	102.500	100.000	136.500
Wilcoxon W	385.500	378.500	376.000	241.500
Z	-1.613	-1.832	-2.055	-.767
Asymp. Sig. (2-tailed)	.107	.067	.040	.443
Exact Sig. [2*(1-tailed Sig.)]	0.107	0.066	0.057	0.448

Since IST-1 and IST-2 trial number satisfied the assumptions for a parametric test, independent t-tests were run for 37 participants (23 females and 14 males) to check for gender differences in both trial numbers at IST-1 and IST-2. For the first phase of the IST, results showed that trial number did not significantly differ between males (M=189.50 SD=57.51) and females (M=156.39, SD=48.33) where $t(35)=1.881$, $p>.05$. However, results showed that IST-2 trial number significantly differed between males (M=152.43 SD=39.35) and females (M=122.78, SD=40.77) where $t(35)=2.173$, $p<.05$. Trial numbers of males were significantly higher than trial numbers of females. Figure 13 shows average values for males and females in IST-1 and IST-2 trial numbers. Summary of the test can be found in Table 7.

Results show that males perform significantly more trials than females during IST-2, even though there is no gender difference in the number of trials during IST-1. Since there is no reward/loss for opening boxes at IST-1, behavior difference in gender is not expected and the result supports this expectation. However, males open more boxes during IST-2 where each opening decreases the reward. Previous studies (van den Bos *et* al. 2013; Overman *et* al., 2006; Stoltenberg & Vandeder, 2010) show that males chose long-term pay-off cases, whereas females prefer larger but infrequent reward choices during reward oriented, investment-based tasks. Females tend to make their choices according to their instincts, whereas males go for more facts and data after analyzing the situation (Agor, 1986; Parik *et* al., 1994). Thus, that shows how our finding about trial numbers was compatible with literature.

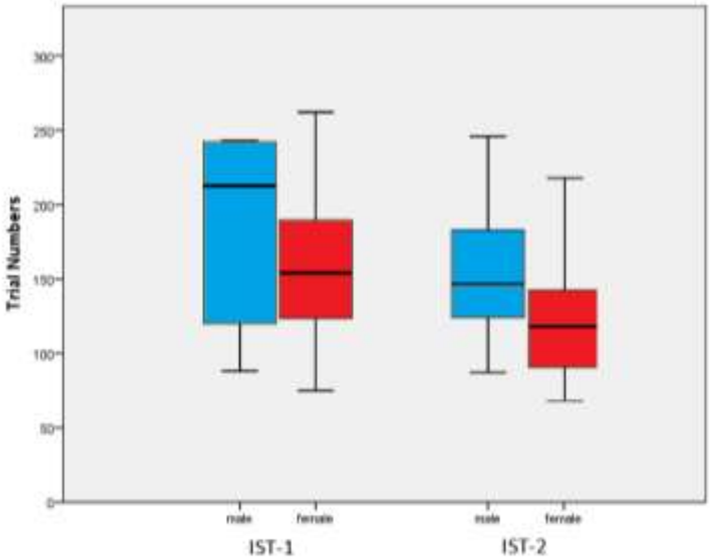

Figure 13: Boxplot of IST Trial Numbers according to Gender

Table 7: Independent t-Test Summary for IST Trial Numbers

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IST1Trial No	Equal variances assumed	1.681	.203	1.881	35	.068	33.109	17.604	-2.628	68.846
	Equal variances not assumed			1.801	23.963	.084	33.109	18.380	-4.828	71.046
IST2Trial No	Equal variances assumed	.080	.779	2.173	35	.037	29.646	13.642	1.951	57.340
	Equal variances not assumed			2.192	28.379	.037	29.646	13.522	1.965	57.327

Since 1-back and 2-back failed the assumptions of normality, a Mann Whitney U-test which is a nonparametric test was conducted for 37 participants (23 females and 14 males) to test for gender difference at each task. Results showed that there was no significant difference between males (M=91.79, SD=14.214) and females (M=92.09, SD=13.031) during the 1-back task where U=160.5, z=-0.19, p>.05. There was no significant difference between males (M=78.86, SD=17.20) and females (M=76.65, SD=23.348) during the 2-back task as well where U=158, z=-0.95, p>.05. Summary of these tests can be found in Table 8. T-tests were conducted for 3-back and 4-back tasks. Results showed that there was no significant difference between males (M=43.36, SD=19.956) and females (M=42.09, SD=19.660) during 3-back task where F=.035, p>.05. However, males 4-back performance (M=41.21, SD=13.735) significantly differed from 4-back females performance (M=26.13, SD=16.926) where F=7.916, p<.05. Figure 14 shows average females and males performance for each back task. Summary of these tests can be found in Table 9.

Table 8: Mann Whitney U-Test Summary for 1-back and 2-back

	back1	back2
Mann-Whitney U	160.500	158.000
Wilcoxon W	436.500	263.000
Z	-.019	-.095
Asymp. Sig. (2-tailed)	.985	.924
Exact Sig. [2*(1-tailed Sig.)]	0.988	0.938

Table 9: t-Test Summary for 3-back and 4-back

Back-3	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	14.041	1	14.041	0.035	0.852
Within Groups	13901	35	397.173		
Total	13915.1	36			
Back-4	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1980.06	1	1980.06	7.916	0.008
Within Groups	8754.97	35	250.142		
Total	10735	36			

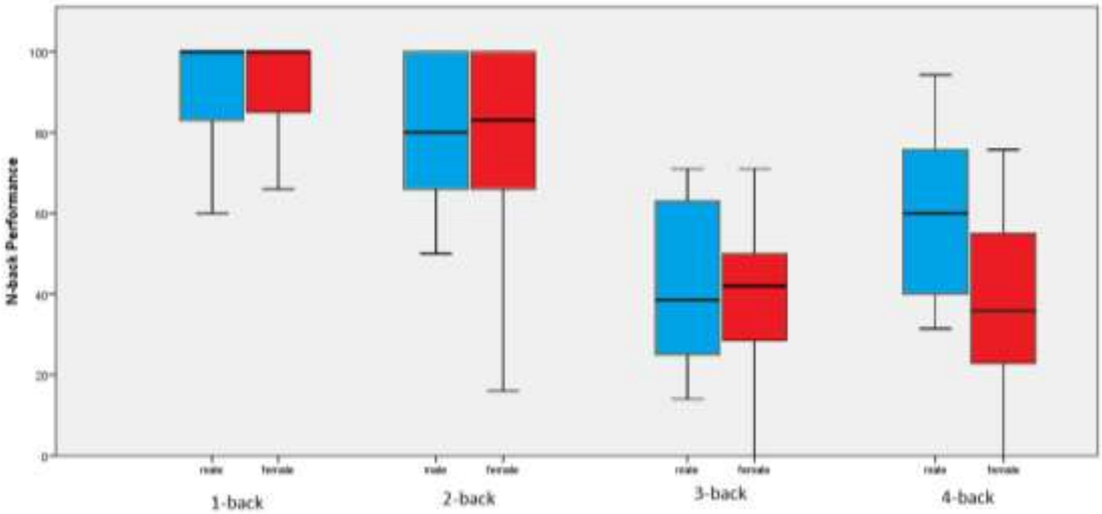

Figure 14: Boxplot of N-back Performances according to Gender

Results suggest that there was no significant difference between gender groups in terms of their performance during 1-back, 2-back and 3-back tests respectively. This finding is consistent with the literature (Li *et* al., 2010; Koch *et* al., 2007; Schmidt *et* al., 2009). However, males are significantly better than females in the 4-back task. Even though there is not a specific study looking for gender differences in 4-back, Voyer *et* al. (1995) show that males had a better performance than females for spatial-based working memory tasks. Our results for the 4-back trial parallel Voyer *et* al.'s finding.

4.3. Task Level Differences in Performance Measures

Maze task consisted of one task while the IST had two phases and the n-back task had 4 phases, therefore level differences were examined between ISTs and n-back tasks.

To investigate performance differences between IST-1 and IST-2 tasks, Wilcoxon Signed Ranks Test was applied for 37 participants (23 females and 14 males) and the results showed that there was a significant difference between performance scores for IST-1 (M=816.22, SD=207.52) and IST-2 (M=399.68, SD=135.57) tasks where z=-5.092, p<.001. Figure 15 shows the boxplot for IST-1 and IST-2 performance scores. Summary of these tests can be found in Table 10.

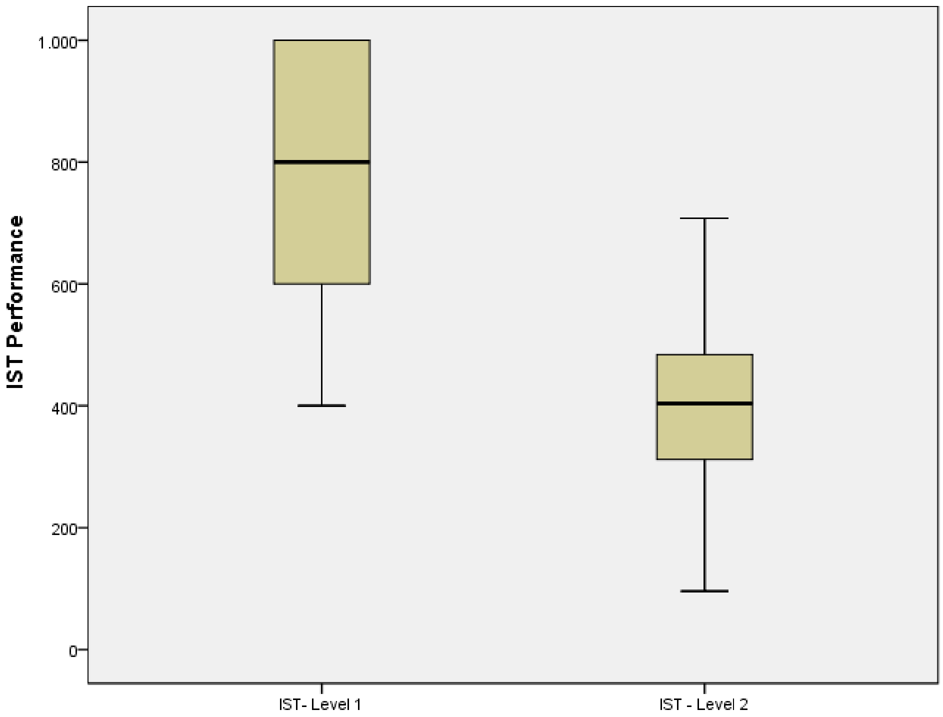

Figure 15: Boxplot of IST Performances according to Task Difficulty

Table 10: Wilcoxon Signed Ranks Test Summary for IST Performances

	IST2_Peformance - IST1_Performance
Z	-5.092
Asymp. Sig. (2-tailed)	.000

Studies on level difference of IST (Solowji *et* al., 2012; Townsheed *et* al. 2006) found out performance in IST-1 was higher than performance in IST-2 and that shows how our finding about level difference was consistent with literature.

Wilcoxon Test was applied for 37 participants (23 females and 14 males) to see whether there was a difference between IST-1 and IST-2 completion time and results showed that there was no significant difference between completion time of IST-1 (M=160.66, SD=61.52) and IST-2 (M=168.92, SD=53.73) tasks, where $z=-1.275$, $p<.05$. Figure 16 shows the boxplots of IST-1 and IST-2 completion times. Summary of these tests can be found in Table 11. There was no significant difference in completion time and that finding was similar with literature (Solowji et al., 2012; Townsheed et al. 2006).

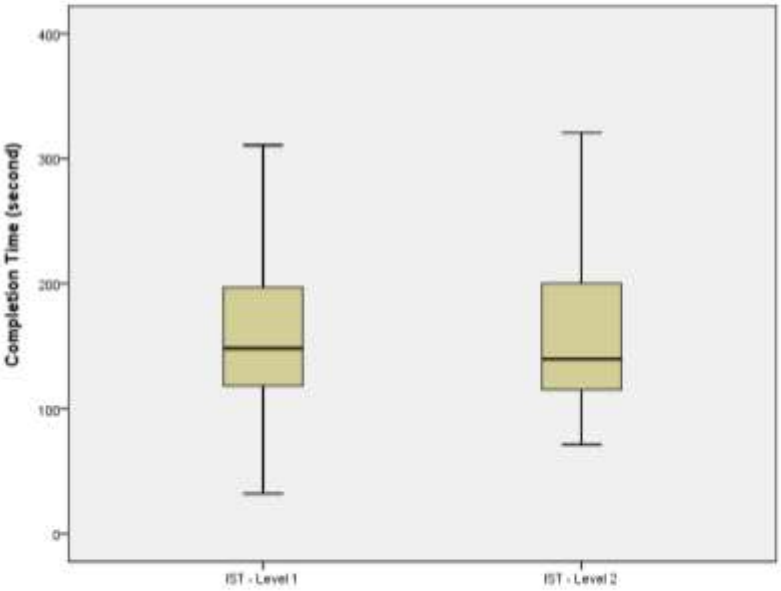

Figure 16: Boxplot of IST Completion Times according to Task Difficulty

Table 11: Wilcoxon Signed Ranks Test Summary for IST Completion Times

	IST2_Time - IST1_Time
Z	-1.275
Asymp. Sig. (2-tailed)	.202

Paired t-test was conducted for 37 participants (23 females and 14 males) to see whether there was a significant difference between IST-1 and IST-2 tasks in terms of number of trials. Results showed that IST-1 trial number (M=168.92, SD=53.730) was significantly higher than IST-2 trial number (M=134.00, SD=42.274) where t(36)=6.06, p<.05. Figure 17 shows average IST-1 and IST-2 trial numbers in a graph. Summary of these tests can be found in Table 12.

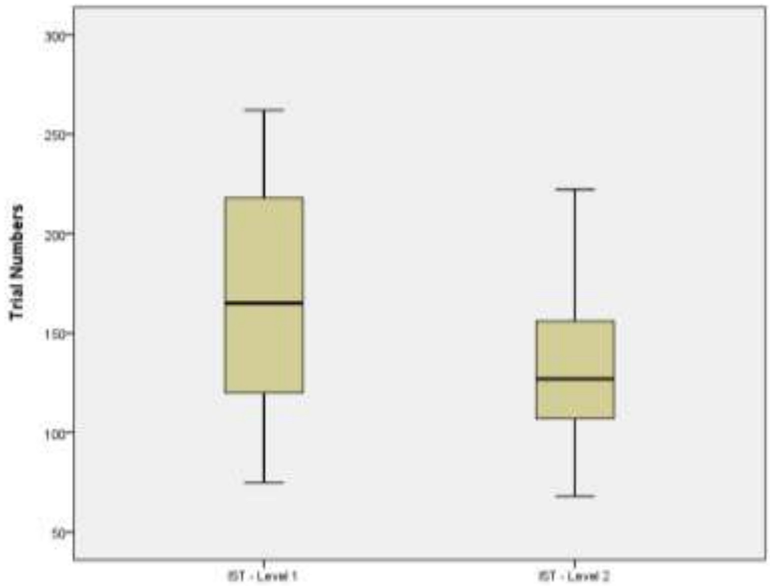

Figure 17: Boxplot of IST Trial Numbers according to Task Difficulty

Table 12: Paired t-Test Summary for IST-1 and IST-2 Trial Numbers

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper			
IST1TrialNo - IST2TrialNo	34.919	35.059	5.764	23.230	46.608	6.058	36	.000

Trial numbers during IST-1 was significantly higher than trial numbers during IST-2 and that result was consistent with literature (Clark *et* al., 2006; Solowji *et* al., 2012). Since there is no reward/loss for trials to open boxes during IST-1, participants opened boxes as many as they wanted. However, penalties for trials prevent increasing numbers of trials at IST-2.

N-back task had four levels and ANOVA with repeated measures was conducted for 37 participants (23 females and 14 males) to see whether performance levels of subjects change according to difficulty level of n-back task. Since $F(2.721, 95.234) = 96.851$, $p<.001$; there was a significant main effect of task level for performance. There was at least one level differing from another level significantly. Summary of the test can be found in Table 13. Figure 18 shows average performances for each level of n-back in a graph.

Existence of performance differences according to task level was also reported in the literature (Ayaz *et* al., 2012; Harvey *et* al., 2005; Blokland *et* al., 2005). Increasing complexity of n-back created a significant decrease in n-back performance in previous studies.

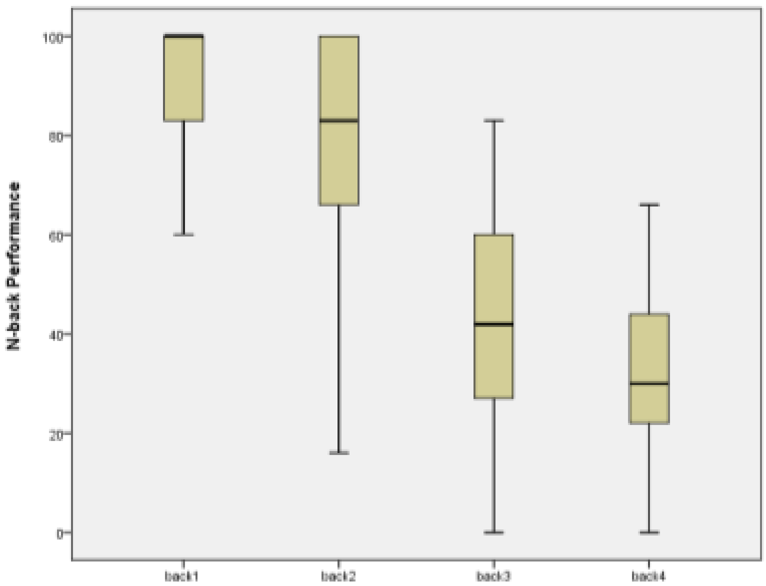

Figure 18: N-back Performance according to Task Difficulty

Table 13: Repeated ANOVA Test Summary for N-back Performances

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
back	Sphericity Assumed	80676.545	3	26892.182	96.851	.000
	Greenhouse-Geisser	80676.545	2.721	29649.782	96.851	.000
	Huynh-Feldt	80676.545	3.000	26892.182	96.851	.000
	Lower-bound	80676.545	1.000	80676.545	96.851	.000
back * Gender	Sphericity Assumed	1311.951	3	437.317	1.575	.200
	Greenhouse-Geisser	1311.951	2.721	482.161	1.575	.204
	Huynh-Feldt	1311.951	3.000	437.317	1.575	.200
	Lower-bound	1311.951	1.000	1311.951	1.575	.218
Error(back)	Sphericity Assumed	29154.793	105	277.665		
	Greenhouse-Geisser	29154.793	95.234	306.137		
	Huynh-Feldt	29154.793	105.000	277.665		
	Lower-bound	29154.793	35.000	832.994		

Pairwise comparisons were applied for 37 participants (23 females and 14 males) to see whether there was a significant difference between any particular pairs of levels and that brings six research questions as follows:

- Do performance levels of subjects change from 1-back to 2-back?
- Do performance levels of subjects change from 1-back to 3-back?
- Do performance levels of subjects change from 1-back to 4-back?
- Do performance levels of subjects change from 2-back to 3-back?
- Do performance levels of subjects change from 2-back to 4-back?
- Do performance levels of subjects change from 3-back to 4-back?

Results show that subjects' performance on 1-back ($M=91.97$, $SD=13.30$) is significantly different from 2-back performance ($M=77.49$, $SD=21.00$) where $p<.05$.

1-back performance is also significantly different from 3-back performance (M=42.57, SD=19.66) and 4-back performance (M=31.84, SD=17.27) where p<.001 for both of them. When we compare 2-back performance with others, it is seen that 2-back performance (M=77.49, SD=21.00) is significantly different from 3-back performance (M=42.57, SD=19.66) and 4-back performance (M=31.84, SD=17.27) where p<.001 for both of them. Lastly, the result show that performance of subjects on 3-back is significantly different from performance on 4-back where p<.05. Statistics of can be seen from Table 14.

Table 14: Pairwise Comparisons for N-back Performances

					95% Confidence Interval for Difference	
(I) back		Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
1-back	2-back	14,182*	3.555	.002	4.240	24.124
	3-back	49,214*	4.153	.000	37.601	60.828
	4-back	58,264*	3.173	.000	49.391	67.137
2-back	1-back	-14,182*	3.555	.002	-24.124	-4.240
	3-back	35,033*	4.377	.000	22.791	47.274
	4-back	44,082*	4.128	.000	32.537	55.628
3-back	1-back	-49,214*	4.153	.000	-60.828	-37.601
	2-back	-35,033*	4.377	.000	-47.274	-22.791
	4-back	9.050	4.423	.290	-3.320	21.419
4-back	1-back	-58,264*	3.173	.000	-67.137	-49.391
	2-back	-44,082*	4.128	.000	-55.628	-32.537
	3-back	-9.050	4.423	.290	-21.419	3.320

4.4. Gender Differences in fNIR Measures

There were 37 participants, but due to excessive noise and motion artifacts, some of the subjects had to be excluded from the fNIR analysis. 19 participants' fNIR measures data out of 37 participants were available to be used after filtering the data. Since fNIR measures satisfy parametric assumptions, parametric tests were conducted to see whether there were differences between gender groups. An independent t-test for the maze task, repeated measures ANOVA for 19 participants (8 females and 11 males) for IST and n-back tasks were conducted to investigate gender differences in oxygenation levels observed in the left and right DLPFC regions. ANOVA for repeated measures for 19 participants (8 females and 11 males) for IST and n-back tasks were conducted to research gender difference for oxygenation level in left and right DLPFC regions. The tests were performed for both oxygenation level in voxel-1 (left DLPFC) and voxel-4 (right DLPFC).

For the maze task, results showed that the mean oxygenation level for left DLPFC did not significantly differ between males (M=0.752, SD=1.118) and females (M=0.384, SD=0.937) where $p>.05$. The mean oxygenation level for right DLPFC did not significantly differ between males (M=0.220, SD=0.0771) and females (M=0.778, SD=0.966) where $p>.05$. Descriptive statistics are presented in Table 15. Summary of the test can be found in Table 16.

Also t-tests were conducted to see any oxygenation level differences between left and right DLPFC regions for only males, but results showed that there is no differences between voxels since $p>.05$. The same test was applied for only females, too. But no oxygenation level differences was observed since $p>.05$.

Table 15: Descriptive Statistics for Oxygenation Levels during Maze Tasks

Gender		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	male	11	0.752	1.118	0.337
	female	8	0.384	0.937	0.331
V4_maze	male	11	0.22	0.771	0.233
	female	8	0.778	0.966	0.341

Table 16: Independent Samples Test for Maze fNIR Measures

		Levene's Test		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
V1_maze	Equal variances assumed	.183	.674	.671	19	.511	.3148	.4689	-.6703	1.2999
	Equal variances not assumed			.692	16.617	.499	.3148	.4550	-.6469	1.2765
V4_maze	Equal variances assumed	3.001	.100	-1.508	19	.149	-.5739	.3805	-1.3733	.2256
	Equal variances not assumed			-1.426	12.313	.179	-.5739	.4024	-1.4482	.3004

Literature (Gron *et* al., 2000) presents that males showed more activation in their left area of their brain whereas females had more activations than males in their right prefrontal cortex during a maze task. However, in this study no gender differences were observed. The fNIR results for gender difference show similarity with performance results of gender difference for maze task, there is no gender differences in oxygenation levels.

The maze task in our study was easier and shorter than the maze tasks typically used in other studies, since there are more cognitive tasks in the experiment. Using longer maze task might have allowed us to see gender differences in brain activation. The number of participants is 19 and expanding the size of participants can give more accurate results. Therefore, using longer maze task and larger sampling may be applied for future studies.

ANOVA for repeated measures was conducted for IST. Results showed that the mean oxygenation level for left DLPFC did not significantly differ between males and females where $F(1,17)=1964$, $p>.05$. If any significant difference was observed, a detail analysis would be applied to see at which level a gender difference was available. However, it was not needed. Descriptive statistics are available in Table 17. Summary of the test can be found in Table 18.

The same test was applied for right DLPFC, too. The mean oxygenation level for right DLPFC did not significantly differ between males and females where $F(1,17)=200$, $p>.05$. Descriptive statistics are available in Table 17. Summary of the test can be found in Table 19.

The finding of no gender difference during a decision-making task does not match with literature. Bolla *et* al. (2004) investigated brain activity during a decision-making task and found out that men showed greater lateralized brain activity to the right hemisphere than females, whereas females showed grater brain activity in the left DLPFC. The reason of mismatch may be the structure of participants. The participants of our study for brain imaging were 11 males and 8 females whose age varied between 21 and 31, while Bolla *et* al. (2004) used 10 males and 10 females whose age varied between 21 and 45. Adding people of various ages to participants might have created diversity which could be the reason of having different results.

Table 17: Descriptive Statistics for Oxygenation Levels during ISTs

Gender		N	Mean	Std. Deviation	Std. Error Mean
V1_ist-1	male	11	0.451	1.088	0.328
	female	8	0.493	0.58	0.205
V4_ist-1	male	11	0.482	0.884	0.267
	female	8	0.634	1.21	0.458
V1_ist-2	male	11	1.324	0.971	0.293
	female	8	0.977	0.891	0.315
V4_ist-2	male	11	1.309	1.118	0.337
	female	8	1.288	0.986	0.349

Table 18: Repeated ANOVA for IST fNIR Measures of Left DLPFC

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
ist_v1	Pillai's Trace	.585	23.949	1	17	.000	.585
	Wilks' Lambda	.415	23.949	1	17	.000	.585
	Hotelling's Trace	1.409	23.949	1	17	.000	.585
	Roy's Largest Root	1.409	23.949	1	17	.000	.585
ist_v1 * Gender	Pillai's Trace	.104	1.964	1	17	.179	.104
	Wilks' Lambda	.896	1.964	1	17	.179	.104
	Hotelling's Trace	.116	1.964	1	17	.179	.104
	Roy's Largest Root	.116	1.964	1	17	.179	.104

Table 19: Repeated ANOVA for IST fNIR Measures of Right DLPFC

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
ist_v4	Pillai's Trace	.462	14.585	1	17	.001	.462
	Wilks' Lambda	.538	14.585	1	17	.001	.462
	Hotelling's Trace	.858	14.585	1	17	.001	.462
	Roy's Largest Root	.858	14.585	1	17	.001	.462
ist_v4 * Gender	Pillai's Trace	.012	0.200	1	17	.661	.012
	Wilks' Lambda	.988	0.200	1	17	.661	.012
	Hotelling's Trace	.012	0.200	1	17	.661	.012
	Roy's Largest Root	.012	0.200	1	17	.661	.012

ANOVA for repeated measures was conducted for n-back tasks. Results showed that the mean oxygenation level for left DLPFC did not significantly differ between males and females where $p>.05$. The same test was also conducted for oxygenation level of right DLPFC during n-back tasks. There was not a significant difference between males and females since $F(3,15)=2936$, $p>.05$. Descriptive statistics are available in Table 20. Summary of the test can be found in Table 21.

The same test was applied for right DLPFC, too. The mean oxygenation level for right DLPFC did not significantly differ between males and females where $F(3,15)=1888$, $p>.05$. Descriptive statistics are available in Table 20. Summary of the test can be found in Table 22.

Table 20: Descriptive Statistics for Oxygenation Levels during N-back Tasks

Gender		N	Mean	Std. Deviation	Std. Error Mean
V1_1back	male	11	0.074	1.288	0.388
	female	8	0.061	0.847	0.299
V4_1back	male	11	0.003	0.996	0.3
	female	8	0.422	0.73	0.258
V1_2back	male	11	0.981	1.314	0.396
	female	8	0.978	1.157	0.409
V4_2back	male	11	0.523	1.364	0.411
	female	8	1.22	1.282	0.453
V1_3back	male	11	1.789	1.268	0.382
	female	8	1.227	1.476	0.522
V4_3back	male	11	1.215	1.584	0.478
	female	8	1.21	1.322	0.467
V1_4back	male	11	2.162	1.992	0.601
	female	8	1.182	1.473	0.521
V4_4back	male	11	1.556	1.947	0.587
	female	8	1.474	1.34	0.474

Table 21: Repeated ANOVA for N-back fNIR Measures of Left DLPFC

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Back_v1	Pillai's Trace	.774	17,136	3.000	15.000	.000	.774
	Wilks' Lambda	.226	17,136	3.000	15.000	.000	.774
	Hotelling's Trace	3.427	17,136	3.000	15.000	.000	.774
	Roy's Largest Root	3.427	17,136	3.000	15.000	.000	.774
Back_v1 * Gender	Pillai's Trace	.370	2,936	3.000	15.000	.067	.370
	Wilks' Lambda	.630	2,936	3.000	15.000	.067	.370
	Hotelling's Trace	.587	2,936	3.000	15.000	.067	.370
	Roy's Largest Root	.587	2,936	3.000	15.000	.067	.370

Table 22: Repeated ANOVA for N-back fNIR Measures of Right DLPFC

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Back_v4	Pillai's Trace	.563	6,443	3.000	15.000	.005	.563
	Wilks' Lambda	.437	6,443	3.000	15.000	.005	.563
	Hotelling's Trace	1.289	6,443	3.000	15.000	.005	.563
	Roy's Largest Root	1.289	6,443	3.000	15.000	.005	.563
Back_v4 * Gender	Pillai's Trace	.274	1,888	3.000	15.000	.175	.274
	Wilks' Lambda	.726	1,888	3.000	15.000	.175	.274
	Hotelling's Trace	.378	1,888	3.000	15.000	.175	.274
	Roy's Largest Root	.378	1,888	3.000	15.000	.175	.274

fNIR results of n-back task show that there is not a significant gender difference. Even though there are previous findings which supported gender differences in brain activation during n-back tasks (Speck *et* al., 2000; Li *et* al., 2010), there are also some studies (Schmidt, 2009; Haut and Barch, 2006) which observed no gender difference in brain. That variation in the literature creates lack of empirical evidence for gender difference of working memory in functional brain organization.

4.5. Task Level Differences in fNIR Measures

Izzetoglu *et* al. (2004) claimed that blood oxygenation level in DLPFC would increase with increasing task difficulty. However, when task became too difficult, a break point was reached and the subject did not pay attention to the task anymore. At that point blood oxygenation level dropped. By considering that hypothesis, a relationship between changes in blood oxygenation and performance level during a cognitive task was investigated.

First, their descriptive statistics are shown in Table 23, Table 24 and Table 25 respectively. Oxygenation levels in voxel-1 (Table 23) show us 4-back has the highest level and 3-back, IST-2, 2-back, maze, IST-1, 1-back are listed from higher to lower respectively. Oxygenation levels in voxel-4 (Table 24) show us almost similar results with left DLPFC. 4-back has the highest level and IST-2, 3-back, 2-

back, IST-1, maze, 1-back are listed from higher to lower respectively.

Table 23: Descriptive Statistics of Oxygenation Levels of Left DLPFC

LEFT	Mean	Std. Deviation	N
V1_maze	0.5731	1.0124	19
V1_ist1	0.4404	0.8739	19
V1_ist2	1.1893	0.9059	19
V1_1back	0.0688	1.0955	19
V1_2back	0.9795	1.2168	19
V1_3back	1.5522	1.3498	19
V1_4back	1.7490	1.8150	19

Table 24: Descriptive Statistics of Oxygenation Levels of Right DLPFC

RIGHT	Mean	Std. Deviation	N
V4_maze	0.4338	0.8612	19
V4_ist1	0.5525	0.9786	19
V4_ist2	1.3113	1.0091	19
V4_1back	0.1792	0.8963	19
V4_2back	0.8162	1.3412	19
V4_3back	1.2128	1.4398	19
V4_4back	1.5218	1.6750	19

The average oxygenation levels of left and right DLPC (Table 25) shows that sequence from higher to lower is as follows: 4-back, 3-back, IST-2, 2-back, maze, IST-1, 1-back. The boxplots of oxygenation levels for left, right and average are shown in Figure 19, 20 and 21 respectively.

Table 25: Descriptive Statistics of Oxygenation Levels for Average of Left and Right DLPFC

AVERAGE	Mean	Std. Deviation	N
Vavg_maze	0.5260	0.7911	19
Vavg_ist1	0.4896	0.8577	19
Vavg_ist2	1.2732	0.9978	19
Vavg_1back	0.1354	0.9231	19
Vavg_2back	0.8979	1.1523	19
Vavg_3back	1.3825	1.3182	19
Vavg_4back	1.6354	1.6525	19

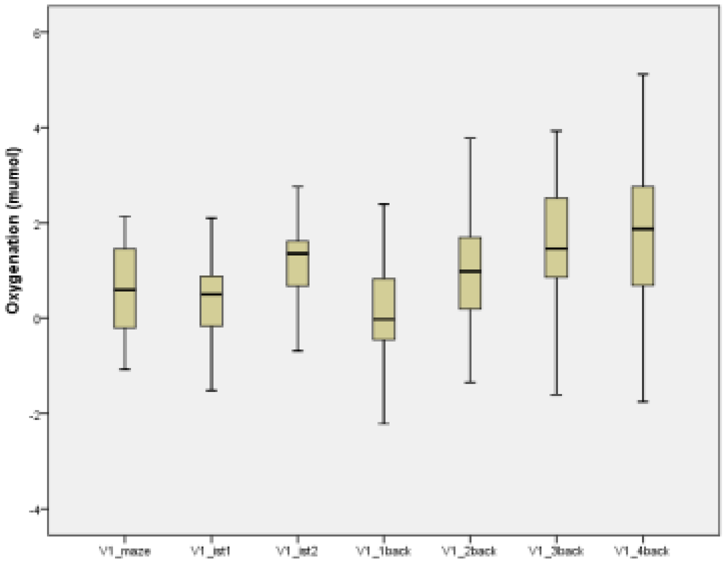

Figure 19: Boxplot of Oxygenation Levels at Left DLPFC

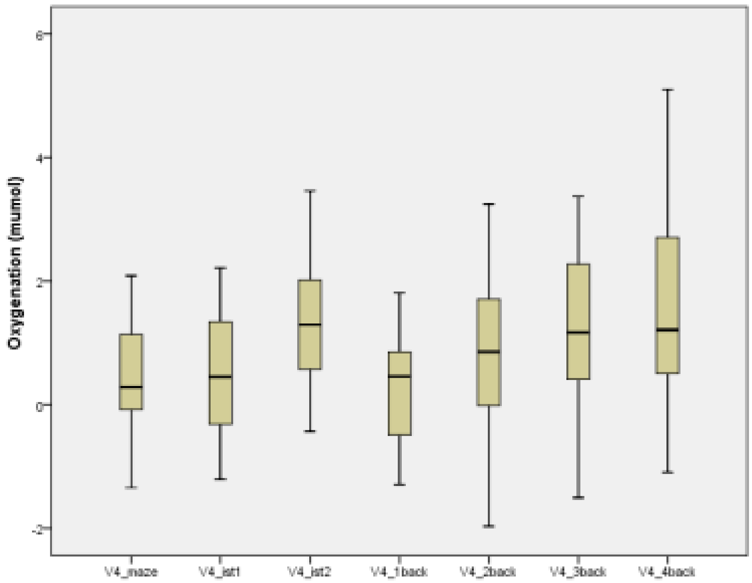

Figure 20: Boxplot of Oxygenation Levels at Right DLPFC

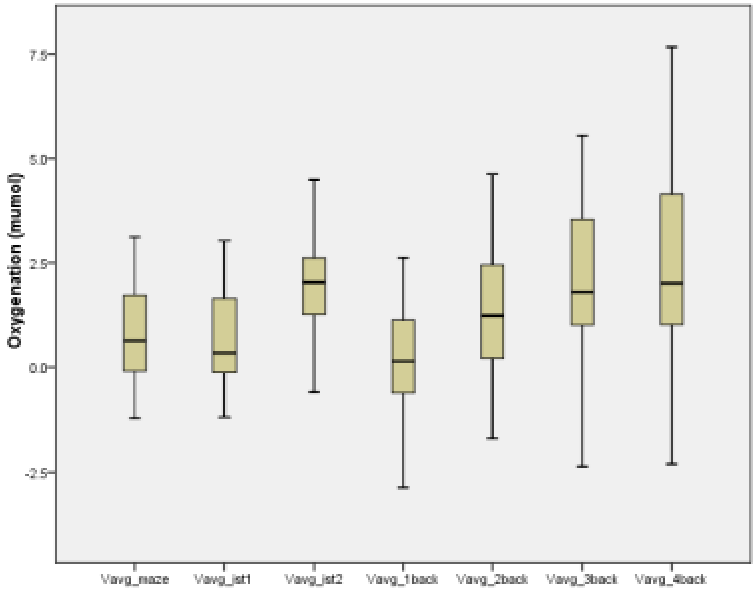

Figure 21: Boxplot of Oxygenation Levels for Average of Left and Right DLPFC

Repeated measures ANOVA with gender as a between-subjects factor and task type (maze, IST-1, IST-2, 1-back, 2-back, 3-back and 4-back) as a within-subjects factor was conducted to compare any cognitive task with other tasks to see any significant difference exists. Repeated measures ANOVA was applied three times for left DLPFC, right DLPFC and average oxygenation level of right and left DLPFC. Results of test showed that there was an overall significant difference between different tasks in oxygenation levels for left DLPFC ($F = 7.439$, $df = 2.403$, $p < .05$), for right DLPFC ($F = 4.848$, $df = 2.6723$, $p < .05$) and for average oxygenation level of right and left DLPFC ($F = 7.221$, $df = 2.516$, $p < .05$). Then, tasks were compared with each other specifically.

First, maze task was compared with other tasks. Results (Table 26) show that, there were significant differences in oxygenation levels at left DLPFC between maze and IST-2 task ($t(19)=-2.221$, $p<.05$), maze and 3-back task ($t(19)=-0.236$, $p<.05$), maze and 4-back task ($t(19)=-0.202$, $p<.05$). However, oxygenation levels at left DLPFC do not differ significantly between maze and IST-1 task ($t(19)=0.497$, $p>.05$), maze and 1-back ($t(19)=1.636$, $p>.05$), maze and 2-back ($t(19)=-1.217$, $p>.05$). Paired t-test for oxygenation level at right DLPFC is shown in Table 27. The oxygenation levels are significantly different between maze and IST-2 tasks ($t(19)=-2.601$, $p<.05$), maze and 3-back tasks ($t(19)=-2.356$, $p<.05$) and maze and 4-back tasks ($t(19)=-2.853$, $p<.05$). There were no significant differences between maze and IST-1 tasks, maze and 1-back tasks and maze and 2-back tasks. The average of oxygenation levels shows same results with left and right DLPFC regions as seen from Table 28. While there were significant differences between maze and IST-2 tasks ($t(19)=-2.601$, $p<.05$), maze and 3-back tasks ($t(19)=-2.356$, $p<.05$), maze and 4-back tasks ($t(19)=-2.853$, $p<.05$); there were not any significant differences between maze and IST-1 ($t(19)= -0.335$, $p>.05$), maze and 1-back ($t(19)=0.997$, $p>.05$), maze and 2-back tasks ($t(19)= -1.155$, $p>.05$).

Table 26: Paired Samples Test to Compare Maze Task with others for Left DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	V1_maze - V1_ist1	.128	1.126	.258	-.415	.671	.497	19	.625
Pair 2	V1_maze - V1_ist2	-.581	1.140	.261	-1.130	-.031	-2.221	19	.039
Pair 3	V1_maze - V1_1back	.528	1.358	.312	-.126	1.183	1.696	19	.107
Pair 4	V1_maze - V1_2back	-.382	1.369	.314	-1.042	.278	-1.217	19	.239
Pair 5	V1_maze - V1_3back	-.955	1.492	.342	-1.674	-.236	-2.790	19	.012
Pair 6	V1_maze - V1_4back	-1.152	1.971	.452	-2.102	-.202	-2.548	19	.020

Table 27: Paired Samples Test to compare Maze Task with others for Right DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	V4_maze - V4_ist1	-.091	1.182	.271	-.661	.479	-.335	19	.742
Pair 2	V4_maze - V4_ist2	-.845	1.417	.325	-1.528	-.162	-2.601	19	.018
Pair 3	V4_maze - V4_1back	.276	1.206	.277	-.305	.857	.997	19	.332
Pair 4	V4_maze - V4_2back	-.361	1.363	.313	-1.018	.296	-1.155	19	.263
Pair 5	V4_maze - V4_3back	-.758	1.402	.322	-1.434	-.082	-2.356	19	.030
Pair 6	V4_maze - V4_4back	-1.067	1.630	.374	-1.853	-.281	-2.853	19	.011

Table 28: Paired Samples Test to compare Maze Task with Others for Average

of Left and Right DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Vavg_maze - Vavg_ist1	.024	1.018	.228	-.453	.500	.105	19	.918
Pair 2	Vavg_maze - Vavg_ist2	-.779	1.238	.277	-1.359	-.200	-2.814	19	.011
Pair 3	Vavg_maze - Vavg_1back	.391	1.175	.270	-.176	.957	1.449	19	.165
Pair 4	Vavg_maze - Vavg_2back	-.372	1.259	.289	-.978	.235	-1.288	19	.214
Pair 5	Vavg_maze - Vavg_3back	-.856	1.343	.308	-1.504	-.209	-2.779	19	.012
Pair 6	Vavg_maze - Vavg_4back	-1.109	1.698	.390	-1.928	-.291	-2.848	19	.011

Second, IST-1 and IST-2 were compared with other tasks to see whether any significant difference exists. Paired t-test for oxygenation level at left DLPFC is shown in Table 29. IST-1 task is significantly different from IST-2 task ($t(19)=-5.047$, $p<.001$), 3-back task ($t(19)=-3.810$, $p<.05$) and 4-back task ($t(19)=-3.298$, $p<.05$) and IST-2 task is significantly different from maze ($t(19)=2.628$, $p<.05$), IST-1 task ($t(19)=-5.047$, $p<.001$) and 1-back task ($t(19)=3.713$, $p<.05$); while there were no significant differences between IST-1 and maze tasks ($t(19)=0.05$, $p>.05$), IST-1 and 1-back tasks ($t(19)=1.582$, $p>.05$), IST-1 and 2-back tasks ($t(19)=-2.038$, $p>.05$), IST-2 and 2-back tasks ($t(19)=0.708$, $p>.05$), IST-2 and 3-back tasks ($t(19)=-1.307$, $p>.05$), IST-2 and 4-back tasks ($t(19)=-1.434$, $p>.05$).

Paired t-tests for right DLPFC is presented in Table 30 and results of right DLPFC is same as left one. IST-1 task is significantly different from IST-2 task ($t(19)=-4.030$, $p<.05$), 3-back task ($t(19)=-2.558$, $p<.05$) and 4-back task ($t(19)=-2.948$, $p<.05$) and IST-2 task is significantly different from maze task ($t(19)=2.601$, $p<.05$), IST-1 task ($t(19)=4.030$, $p<.05$) and 1-back task ($t(19)=3.763$, $p<.05$); while there are no significantly differences between IST-1 and maze tasks ($t(19)=0.05$, $p>.05$), IST-1 and 1-back tasks ($t(19)=1.754$, $p>.05$), IST-1 and 2-back tasks ($t(19)=-0.983$, $p>.05$), IST-2 and 2-back tasks ($t(19)=-0.983$, $p>.05$), IST-2 and 3-back tasks ($t(19)=0.235$, $p>.05$), IST-2 and 4-back tasks ($t(19)=-0.498$, $p>.05$).

Since paired t-test for left and right DLPFC regions show same results their average will show the same. Despite knowing that, paired t-test was also conducted for average oxygenation levels as seen from Table 31.

Table 29: Paired Samples Test to Compare IST with others for Left DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	V1_ist1 - V1_maze	.014	1.210	.278	-.569	.597	.050	19	.961
Pair 2	V1_ist1 - V1_ist2	-.709	.612	.140	-1.004	-.414	-5.047	19	.000
Pair 3	V1_ist1 - V1_1back	.400	1.102	.253	-.131	.931	1.582	19	.131
Pair 4	V1_ist1 - V1_2back	-.511	1.093	.251	-1.037	.016	-2.038	19	.057
Pair 5	V1_ist1 - V1_3back	-1.083	1.240	.284	-1.681	-.486	-3.810	19	.001
Pair 6	V1_ist1 - V1_4back	-1.280	1.692	.388	-2.096	-.465	-3.298	19	.004
Pair 7	V1_ist2 - V1_maze	.723	1.199	.275	.145	1.301	2.628	19	.017
Pair 8	V1_ist2 - V1_ist1	.709	.612	.140	.414	1.004	5.047	19	.000
Pair 9	V1_ist2 - V1_1back	1.109	1.302	.299	.481	1.736	3.713	19	.002
Pair 10	V1_ist2 - V1_2back	.198	1.220	.280	-.390	.786	.708	19	.488
Pair 11	V1_ist2 - V1_3back	-.374	1.249	.286	-.976	.227	-1.307	19	.208
Pair 12	V1_ist2 - V1_4back	-.571	1.737	.398	-1.409	.266	-1.434	19	.169

Table 30: Paired Samples Test to Compare IST with others for Right DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	V4_ist1 - V4_maze	.014	1.210	.278	-.569	.597	.050	18	.961
Pair 2	V4_ist1 - V4_ist2	-.755	.816	.187	-1.148	-.361	-4.030	18	.001
Pair 3	V4_ist1 - V4_1back	.367	.911	.209	-.073	.806	1.754	18	.096
Pair 4	V4_ist1 - V4_2back	-.270	1.199	.275	-.848	.307	-.983	18	.338
Pair 5	V4_ist1 - V4_3back	-.667	1.136	.261	-1.215	-.119	-2.558	18	.020
Pair 6	V4_ist1 - V4_4back	-.976	1.443	.331	-1.672	-.280	-2.948	18	.009
Pair 7	V4_ist2 - V4_maze	.845	1.417	.325	.162	1.528	2.601	18	.018
Pair 8	V4_ist2 - V4_ist1	.755	.816	.187	.361	1.148	4.030	18	.001
Pair 9	V4_ist2 - V4_1back	1.121	1.299	.298	.495	1.747	3.763	18	.001
Pair 10	V4_ist1 - V4_2back	-.270	1.199	.275	-.848	.307	-.983	18	.338
Pair 11	V4_ist2 - V4_3back	.088	1.626	.373	-.696	.871	.235	18	.817
Pair 12	V4_ist2 - V4_4back	-.221	1.939	.445	-1.156	.713	-.498	18	.625

Table 31: Paired Samples Test to Compare IST with others for Average of Left and Right DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Vavg_ist1 - Vavg_maze	-.024	1.018	.228	-.500	.453	-.105	19	.918
Pair 2	Vavg_ist1 - Vavg_ist2	-.803	.621	.139	-1.094	-.512	-5.780	19	.000
Pair 3	Vavg_ist1 - Vavg_1back	.354	.947	.217	-.102	.811	1.631	18	.120
Pair 4	Vavg_ist1 - Vavg_2back	-.408	1.026	.235	-.903	.086	-1.734	18	.100
Pair 5	Vavg_ist1 - Vavg_3back	-.893	1.118	.257	-1.432	-.354	-3.480	18	.003
Pair 6	Vavg_ist1 - Vavg_4back	-1.146	1.469	.337	-1.854	-.438	-3.399	18	.003
Pair 7	Vavg_ist2 - Vavg_maze	.779	1.238	.277	.200	1.359	2.814	19	.011
Pair 8	Vavg_ist2 - Vavg_ist1	.803	.621	.139	.512	1.094	5.780	19	.000
Pair 9	Vavg_ist2 - Vavg_1back	1.138	1.283	.294	.519	1.756	3.865	18	.001
Pair 10	Vavg_ist2 - Vavg_2back	.375	1.312	.301	-.257	1.008	1.247	18	.228
Pair 11	Vavg_ist2 - Vavg_3back	-.109	1.356	.311	-.763	.544	-.351	18	.730
Pair 12	Vavg_ist2 - Vavg_4back	-.362	1.713	.393	-1.188	.464	-.922	18	.369

Lastly, n-back tasks were compared with other tasks. Paired t-test for oxygenation level at left DLPFC is shown in Table 32. 1-back task is significantly different from IST-2 task ($t(19)=-3.713$, $p<.05$), 2-back task ($t(19)=-5.909$, $p<.001$), 3-back task (t(19)=-7.110, p<.001) and 4-back task (t(19)=-6.627, p<.001); 2-back task is significantly different from 1-back (t(19)=-5.909, p<.001), 3-back (t(19)=5.062, p<.001) and 4-back (t(19)=-3.297, p<.05) tasks; 3-back is significantly different from IST-1 task (t(19)=3.810, p<.05), maze (t(19)=2.790, p<.05), 1-back (t(19)=-7.110, p<.001) and 2-back tasks (t(19)=5.062, p<.001); 4-back is significantly different from IST-1 (t(19)=3.298, p<.05), maze (t(19)=2.548, p<.05), 1-back (t(19)=-6.627, p<.001) and 2-back (t(19)=-3.297, p<.05) tasks.

Paired t-test for oxygenation level of right DLPFC was conducted and is presented in Table 33. 1-back task is significantly different from IST-2 (t(19)=-3.763, p<.05), 2-back (t(19)=-3.587, p<.05), 3-back (t(19)=,-4.706 p<.001) and 4-back (t(19)=-4.712, p<.001) tasks; 2-back task is significantly different from 1-back (t(19)=-2.457, p<.05), 3-back (t(19)=-4.706, p<.05) and 4-back (t(19)=-2.894, p<.05) tasks; 3-back is significantly different from IST-1 task (t(19)=2.558, p<.05), maze (t(19)=2.356, p<.05), 1-back (t(19)=-4.706, p<.001), 2-back (t(19)=-2.457, p<.05) and 4-back (t(19)=-2.178, p<.05) tasks; 4-back is significantly different from IST-1 (t(19)=2.948, p<.05), maze (t(19)=2.853, p<.05), 1-back (t(19)=-4.712, p<.001), 2-back (t(19)=-2.894, p<.05) and 3-back (t(19)=-2.178, p<.05) tasks.

Even though paired t-tests for right and left DLPFC are similar, they are not the same. Therefore, another paired t-test was conducted for their average oxygenation levels (Table 34). Results show that 1-back task is significantly different from IST-2 (t(19)=,-1.631 p<.05), 2-back (t(19)=-5.167, p<.05), 3-back (t(19)=-6.532, p<.05) and 4-back (t(19)=-6.367, p<.05) tasks; 2-back task is significantly different from 1-back (t(19)=-5.167, p<.05), 3-back (t(19)=-4.046, p<.05) and 4-back (t(19)=3.465, p<.05) tasks; 3-back is significantly different from IST-1 task (t(19)=3.480, p<.05), maze (t(19)=2.779, p<.05), 1-back (t(19)=6.532, p<.05) and 2-back (t(19)=4.046, p<.05) tasks; 4-back is significantly different from IST-1 (t(19)=3.399, p<.05), maze (t(19)=2.848, p<.05), 1-back (t(19)=6.367, p<.05) and 2-back (t(19)=3.465, p<.05) tasks .

Table 32: Paired Samples Test for compare N-back Tasks with Others for Left DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper			
Pair 1	V1_1back - V1_ist1	-.400	1.102	.253	-.931	.131	-1.582	19	.131
Pair 2	V1_1back - V1_ist2	-1.109	1.302	.299	-1.736	-.481	-3.713	19	.002
Pair 3	V1_1back - V1_maze	-.528	1.358	.312	-1.183	.126	-1.696	19	.107
Pair 4	V1_1back - V1_2back	-.911	.672	.154	-1.235	-.587	-5.909	19	.000
Pair 5	V1_1back - V1_3back	-1.483	.909	.209	-1.922	-1.045	-7.110	19	.000
Pair 6	V1_1back - V1_4back	-1.680	1.105	.254	-2.213	-1.148	-6.627	19	.000
Pair 7	V1_2back - V1_ist1	.511	1.093	.251	-.016	1.037	2.038	19	.057
Pair 8	V1_2back - V1_ist2	-.198	1.220	.280	-.786	.390	-.708	19	.488
Pair 9	V1_2back - V1_maze	.382	1.369	.314	-.278	1.042	1.217	19	.239
Pair 10	V1_2back - V1_1back	.911	.672	.154	.587	1.235	5.909	19	.000
Pair 11	V1_2back - V1_3back	-.573	.493	.113	-.810	-.335	-5.062	19	.000
Pair 12	V1_2back - V1_4back	-.769	1.017	.233	-1.260	-.279	-3.297	19	.004
Pair 13	V1_3back - V1_ist1	1.083	1.240	.284	.486	1.681	3.810	19	.001
Pair 14	V1_3back - V1_ist2	.374	1.249	.286	-.227	.976	1.307	19	.208
Pair 15	V1_3back - V1_maze	.955	1.492	.342	.236	1.674	2.790	19	.012
Pair 16	V1_3back - V4_1back	1.373	1.101	.252	.842	1.903	5.438	19	.000
Pair 17	V1_3back - V1_2back	.573	.493	.113	.335	.810	5.062	19	.000
Pair 18	V1_3back - V1_4back	-.197	.837	.192	-.600	.206	-1.026	19	.319
Pair 19	V1_4back - V1_ist1	1.280	1.692	.388	.465	2.096	3.298	19	.004
Pair 20	V1_4back - V1_ist2	.571	1.737	.398	-.266	1.409	1.434	19	.169
Pair 21	V1_4back - V1_maze	1.152	1.971	.452	.202	2.102	2.548	19	.020
Pair 22	V1_4back - V1_1back	1.680	1.105	.254	1.148	2.213	6.627	19	.000
Pair 23	V1_4back - V1_2back	.769	1.017	.233	.279	1.260	3.297	19	.004
Pair 24	V1_4back - V1_3back	.197	.837	.192	-.206	.600	1.026	19	.319

Table 33: Paired Samples Test for compare N-back Tasks with Others for Right DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	V4_1back - V4_ist1	-.367	.911	.209	-.806	.073	-1.754	19	.096
Pair 2	V4_1back - V4_ist2	-1.121	1.299	.298	-1.747	-.495	-3.763	19	.001
Pair 3	V4_1back - V4_maze	-.276	1.206	.277	-.857	.305	-.997	19	.332
Pair 4	V4_1back - V4_2back	-.637	.774	.178	-1.010	-.264	-3.587	19	.002
Pair 5	V4_1back - V4_3back	-1.034	.957	.220	-1.495	-.572	-4.706	19	.000
Pair 6	V4_1back - V4_4back	-1.343	1.242	.285	-1.941	-.744	-4.712	19	.000
Pair 7	V4_2back - V4_ist1	.270	1.199	.275	-.307	.848	.983	19	.338
Pair 8	V4_2back - V4_ist2	-.484	1.589	.365	-1.250	.282	-1.328	19	.201
Pair 9	V4_2back - V4_maze	.361	1.363	.313	-.296	1.018	1.155	19	.263
Pair 10	V4_2back - V4_1back	.637	.774	.178	.264	1.010	3.587	19	.002
Pair 11	V4_2back - V4_3back	-.397	.703	.161	-.736	-.057	-2.457	19	.024
Pair 12	V4_2back - V4_4back	-.706	1.063	.244	-1.218	-.193	-2.894	19	.010
Pair 13	V4_3back - V4_ist1	.667	1.136	.261	.119	1.215	2.558	19	.020
Pair 14	V4_3back - V4_ist2	-.088	1.626	.373	-.871	.696	-.235	19	.817
Pair 15	V4_3back - V4_maze	.758	1.402	.322	.082	1.434	2.356	19	.030
Pair 16	V4_3back - V4_1back	1.034	.957	.220	.572	1.495	4.706	19	.000
Pair 17	V4_3back - V4_2back	.397	.703	.161	.057	.736	2.457	19	.024
Pair 18	V4_3back - V4_4back	-.309	.619	.142	-.607	-.011	-2.178	19	.043
Pair 19	V4_4back - V4_ist1	.976	1.443	.331	.280	1.672	2.948	19	.009
Pair 20	V4_4back - V4_ist2	.221	1.939	.445	-.713	1.156	.498	19	.625
Pair 21	V4_4back - V4_maze	1.067	1.630	.374	.281	1.853	2.853	19	.011
Pair 22	V4_4back - V4_1back	1.343	1.242	.285	.744	1.941	4.712	19	.000
Pair 23	V4_4back - V4_2back	.706	1.063	.244	.193	1.218	2.894	19	.010
Pair 24	V4_4back - V4_3back	.309	.619	.142	.011	.607	2.178	19	.043

Table 34: Paired Samples Test for compare N-back Tasks with Others for Average of Left and Right DLPFC

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Vavg_1back - Vavg_maze	-.391	1.175	.270	-.957	.176	-1.449	19	.165
Pair 2	Vavg_1back - Vavg_ist1	-.354	.947	.217	-.811	.102	-1.631	19	.120
Pair 3	Vavg_1back - Vavg_ist2	-1.138	1.283	.294	-1.756	-.519	-3.865	19	.001
Pair 4	Vavg_1back - Vavg_2back	-.763	.643	.148	-1.073	-.452	-5.167	19	.000
Pair 5	Vavg_1back - Vavg_3back	-1.247	.832	.191	-1.648	-.846	-6.532	19	.000
Pair 6	Vavg_1back - Vavg_4back	-1.500	1.027	.236	-1.995	-1.005	-6.367	19	.000
Pair 7	Vavg_2back - Vavg_maze	.372	1.259	.289	-.235	.978	1.288	19	.214
Pair 8	Vavg_2back - Vavg_ist1	.408	1.026	.235	-.086	.903	1.734	19	.100
Pair 9	Vavg_2back - Vavg_ist2	-.375	1.312	.301	-1.008	.257	-1.247	19	.228
Pair 10	Vavg_2back - Vavg_1back	.763	.643	.148	.452	1.073	5.167	19	.000
Pair 11	Vavg_2back - Vavg_3back	-.485	.522	.120	-.736	-.233	-4.046	19	.001
Pair 12	Vavg_2back - Vavg_4back	-.738	.928	.213	-1.185	-.290	-3.465	19	.003
Pair 13	Vavg_3back - Vavg_maze	.856	1.343	.308	.209	1.504	2.779	19	.012
Pair 14	Vavg_3back - Vavg_ist1	.893	1.118	.257	.354	1.432	3.480	19	.003
Pair 15	Vavg_3back - Vavg_ist2	.109	1.356	.311	-.544	.763	.351	19	.730
Pair 16	Vavg_3back - Vavg_1back	1.247	.832	.191	.846	1.648	6.532	19	.000
Pair 17	Vavg_3back - Vavg_2back	.485	.522	.120	.233	.736	4.046	19	.001
Pair 18	Vavg_3back - Vavg_4back	-.253	.648	.149	-.565	.060	-1.700	19	.106
Pair 19	Vavg_4back - Vavg_maze	1.109	1.698	.390	.291	1.928	2.848	19	.011
Pair 20	Vavg_4back - Vavg_ist1	1.146	1.469	.337	.438	1.854	3.399	19	.003
Pair 21	Vavg_4back - Vavg_ist2	.362	1.713	.393	-.464	1.188	.922	19	.369
Pair 22	Vavg_4back - Vavg_1back	1.500	1.027	.236	1.005	1.995	6.367	19	.000
Pair 23	Vavg_4back - Vavg_2back	.738	.928	.213	.290	1.185	3.465	19	.003
Pair 24	Vavg_4back - Vavg_3back	.253	.648	.149	-.060	.565	1.700	19	.106

To sum up, from low oxygenation level to highest, tasks can be ranked as follows: 1-back, IST-1, maze, 2-back, IST-2, 3-back, 4-back in voxel-1; 1-back, maze, IST-1, 2-back, IST-2, 3-back, 4-back in voxel-4; and lastly 1-back, IST-1, maze, 2-back, IST-2, 3-back, 4-back at average oxygenation levels of left and right DLPFC regions. These rankings and significant differences give an idea about task difficulty of different kind of cognitive tasks. For example, when IST-1 is compared with n-back tasks difficulty of the task is between 1-back and 2-back according to rankings and paired t-test show actually there was no significant differences in their blood oxygenation levels. Therefore, we can conclude that the difficulty of the task is similar with 1-back and 2-back.

Blood oxygenation level provide assessing mental workload of participants, so comparing these oxygenation levels during cognitive tasks gives information about difficulties of tasks.

4.6. The Relationship of fNIR Measures

fNIR data was collected from all 37 participants, but physical problems (different shapes of foreheads) raise difficulties in collecting fNIR data from all participants from all voxels. Location of voxels is shown in Figure 22. There are 2 detectors on each fNIR sensor and that means there are 4 channels provide hemoglobin and oxy-hemoglobin level data. After filtering fNIR data by FIR and SMAR, 19 participants' data of voxel-1 and voxel-4 were available to be used.

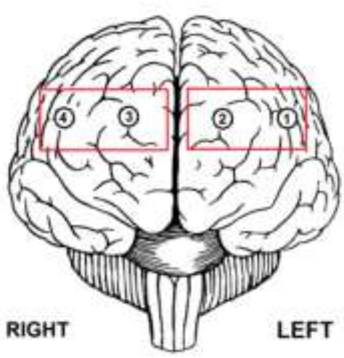

Figure 22: Location of Voxels at Prefrontal Cortex

Since voxel-1 is gathering information about situation of left DLPFC and voxel-4 is gathering information about situation of right DLPFC, first research question of fNIR data about whether there is a difference between changes in oxygenation levels of left DLPFC and right DLPFC during a cognitive task. Paired t-test were applied for 19 participants (8 females and 11 males) for each cognitive tasks and results show that there was no significant differences between changes in oxygenation levels of left DLPFC and right DLPFC during maze task, IST-1 and IST-2 tasks, and all n-back tasks since $p>.05$. Means and standard deviations of oxygenation levels can be seen from Table 35.The result of paired t-test is shown in Table 36.

Table 35: Descriptive Statistics for Oxygenation Levels at Left and Right DLPFC

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	V1_maze (Left)	.60	19	1.034	0.237
	V4_maze (Right)	.45	19	0.879	0.202
Pair 2	V1_ist1 (Left)	.47	19	0.888	0.204
	V4_ist1 (Right)	.55	19	1.005	0.231
Pair 3	V1_ist2 (Left)	1.18	19	0.929	0.213
	V4_ist2 (Right)	1.30	19	1.036	0.238
Pair 4	V1_1back (Left)	.07	19	1.096	0.251
	V4_1back (Right)	.18	19	0.896	0.206
Pair 5	V1_2back (Left)	.98	19	1.217	0.279
	V4_2back (Right)	.82	19	1.341	0.308
Pair 6	V1_3back (Left)	1.55	19	1.350	0.310
	V4_3back (Right)	1.21	19	1.440	0.330
Pair 7	V1_4back (Left)	1.75	19	1.815	0.416
	V4_4back (Right)	1.52	19	1.675	0.384

Table 36: Paired t-Test for Oxygenation Levels at Left and Right DLPFC

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	V1_maze - V4_maze	.142	1.088	.250	-.382	.666	.570	19	.576
Pair 2	V1_ist1 - V4_ist1	-.077	.906	.208	-.514	.360	-.370	19	.715
Pair 3	V1_ist2 - V4_ist2	-.123	.778	.178	-.498	.252	-.687	19	.501
Pair 4	V1_1back - V4_1back	-.110	.856	.196	-.523	.302	-.562	19	.581
Pair 5	V1_2back - V4_2back	.163	1.117	.256	-.375	.702	.637	19	.532
Pair 6	V1_3back - V4_3back	.339	.916	.210	-.102	.781	1.614	19	.124
Pair 7	V1_4back - V4_4back	.227	1.130	.259	-.318	.772	.876	19	.392

4.7. Correlation between fNIR and Performance Measures

Another research question is about the relationship between fNIR and performance measures. Therefore, correlations between measures were investigated for each task. Pearson correlation was applied. Maze path length and maze time are two performance measures for maze task and the oxygenation level at right DLPFC and the oxygenation level at left DLPFC are two fNIR measures for maze. Four different measures were compared by Pearson correlation and results are shown in Table 37.

Since $p<.05$, there is a significant linear relationship between maze path length and maze time. The Pearson correlation is 0.816. This means there is a strong positive relationship between maze path length and time. Since other p values are greater than .05, it cannot be said that there is any correlation between them.

Table 37: Correlations between Maze Measures

		MazePath	MazeTime	V1_maze	V4_maze
MazePath	Pearson	1	0.816**	0.078	-0.113
	Sig. (2-tailed)		0	0.752	0.646
	N	19	19	19	19
MazeTime	Pearson	0.816**	1	-0.064	0.03
	Sig. (2-tailed)	0		0.796	0.903
	N	19	19	19	19
V1_maze	Pearson	0.078	-0.064	1	0.363
	Sig. (2-tailed)	0.752	0.796		0.127
	N	19	19	19	19
V4_maze	Pearson	-0.113	0.03	0.363	1
	Sig. (2-tailed)	0.646	0.903	0.127	
	N	19	19	19	19

Performance, trial number and completion time are used as performance measures and the oxygenation level at right DLPFC and the oxygenation level at left DLPFC are two fNIR measures for ISTs. The result of Pearson correlation for IST-1 is presented in Table 38. There are significant relationships between performance and trial number, performance and time, trial number and time, the oxygenation level at right DLPFC and the oxygenation level at left DLPFC since p<.05. Pearson correlation between IST-1 performance and IST-1 Trial number is 0.717, so there is a strong positive correlation between them. Pearson correlation between IST-1 performance and IST-1 time is 0.508, so there is a strong positive correlation between them. Pearson correlation between IST-1 trial number and IST-1 time is 0.768, so there is a strong positive correlation between them. Pearson correlation between the oxygenation level in voxel-1 and voxel-4 is 0.548, so there is a strong positive correlation between them (Table 38).

Table 38: Correlations between IST-1 Measures

		IST1-Per	IST1-TrialNo	IST1-Time	V1_ist1	V4_ist1
IST1-Per	Pearson	1	0.717**	0.508*	-0.184	-0.200
	Sig. (2-tailed)		0.001	0.026	0.452	0.411
	N	19	19	19	19	19
IST1-TrialNo	Pearson	0.717**	1	0.768**	0.004	-0.314
	Sig. (2-tailed)	0.001		0.000	0.986	0.190
	N	19	19	19	19	19
IST1-Time	Pearson	0.508*	0.768**	1	-0.183	-0.252
	Sig. (2-tailed)	0.026	0.000		0.453	0.298
	N	19	19	19	19	19
V1_ist1	Pearson	-0.184	0.004	-0.183	1	0.548*
	Sig. (2-tailed)	0.452	0.986	0.453		0.015
	N	19	19	19	19	19
V4_ist1	Pearson	-0.200	-0.314	-0.252	0.548*	1
	Sig. (2-tailed)	0.411	0.190	0.298	0.015	
	N	19	19	19	19	19

The result of Pearson correlation for IST-2 to investigate relationship of 5 measures is presented in Table 39. There are significant relationships between performance and trial number, trial number and time, the oxygenation level at right DLPFC and the oxygenation level at left DLPFC since $p<.05$. Pearson correlation between IST-2 performance and IST-2 Trial number is -0.499, so there is a negative moderate correlation between them. Pearson correlation between IST-2 trial number and IST-2 time is 0.809, so there is a strong positive correlation between them. Pearson correlation between the oxygenation level in voxel-1 and voxel-4 is 0.692, so there is a strong positive correlation between them (Table 39).

Table 39: Correlations between IST-2 Measures

		IST2-Per	IST2-TrialNo	IST2-Time	V1_ist2	V4_ist2
IST2-Per	Pearson	1	-0.499	-0.304	-0.386	-0.280
	Sig. (2-tailed)		0.030	0.206	0.103	0.245
	N	19	19	19	19	19
IST2-TrialNo	Pearson	-0.499	1	0.809**	-0.029	-0.094
	Sig. (2-tailed)	0.030		0.000	0.908	0.703
	N	19	19	19	19	19
IST2-Time	Pearson	-0.304	0.809**	1	0.042	0.008
	Sig. (2-tailed)	0.206	0.000		0.865	0.975
	N	19	19	19	19	19
V1_ist2	Pearson	-0.386	-0.029	0.042	1	0.692**
	Sig. (2-tailed)	0.103	0.908	0.865		0.001
	N	19	19	19	19	19
V4_ist2	Pearson	-0.280	-0.094	0.008	0.692**	1
	Sig. (2-tailed)	0.245	0.703	0.975	0.001	
	N	19	19	19	19	19

Performance is one and only performance measure and the oxygenation level at right DLPFC and the oxygenation level at left DLPFC are two fNIR measures for n-back tasks. Four different Pearson correlations were conducted to see relationship for each task. As seen from Table 40, there is a significant relationship between the oxygenation level at right DLPFC and the oxygenation level at left DLPFC since $p<.05$ but performance does not have any relationship with fNIR measure at .05 level. Pearson correlation between the oxygenation level in voxel-1 and voxel-4 is 0.657, so there is a strong positive correlation between them.

Pearson correlation for 2-back task can be seen from Table 41 and results show that performance as performance measure does not have a significant relationship with fNIR measures while there is a significant relationship the oxygenation level at right DLPFC and the oxygenation level at left DLPFC since $p<.05$. Pearson correlation between the oxygenation level in voxel-1 and voxel-4 is 0.647, so there is a strong positive correlation between them.

Table 40: Correlations between 1-back Measures

		back1	V1_1back	V4_1back
back1	Pearson	1	0.084	0.022
	Sig. (2-tailed)		0.732	0.928
	N	19	19	19
V1_1back	Pearson	0.084	1	0.647**
	Sig. (2-tailed)	0.732		0.003
	N	19	19	19
V4_1back	Pearson	0.022	0.647**	1
	Sig. (2-tailed)	0.928	0.003	
	N	19	19	19

Table 41: Correlations between 2-back Measures

		back2	V1_2back	V4_2back
back2	Pearson	1	0.048	0.128
	Sig. (2-tailed)		0.846	0.603
	N	19	19	19
V1_2back	Pearson	0.048	1	0.623**
	Sig. (2-tailed)	0.846		0.004
	N	19	19	19
V4_2back	Pearson	0.128	0.623**	1
	Sig. (2-tailed)	0.603	0.004	
	N	19	19	19

Pearson correlation for 3-back task can be seen from Table 42 and results show that performance as performance measure does not have a significant relationship with fNIR measures while there is a significant relationship the oxygenation level at right DLPFC and the oxygenation level at left DLPFC since $p<.05$. Pearson correlation between the oxygenation level in voxel-1 and voxel-4 is 0.786, so there is a strong positive correlation between them.

Table 42: Correlations between 3-back Measures

		back3	V1_3back	V4_3back
back3	Pearson	1	-0.051	0.137
	Sig. (2-tailed)		0.834	0.577
	N	19	19	19
V1_3back	Pearson	-0.051	1	0.786**
	Sig. (2-tailed)	0.834		0.000
	N	19	19	19
V4_3back	Pearson	0.137	0.786**	1
	Sig. (2-tailed)	0.577	0.000	
	N	19	19	19

Pearson correlation for 4-back task can be seen from Table 43 and results show that performance as a performance measure does not have a significant relationship with fNIR measures while there is a significant relationship between the oxygenation level at right DLPFC and the oxygenation level at left DLPFC since $p<.05$. Pearson correlation between the oxygenation level in voxel-1 and voxel-4 is 0.534, so there is a strong positive correlation between them.

Table 43: Correlations between 4-back Measures

		back4	V1_4back	V4_4back
back4	Pearson	1	0.397	0.152
	Sig. (2-tailed)		0.093	0.534
	N	19	19	19
V1_4back	Pearson	0.397	1	0.793**
	Sig. (2-tailed)	0.093		0.000
	N	19	19	19
V4_4back	Pearson	0.152	0.793**	1
	Sig. (2-tailed)	0.534	0.000	
	N	19	19	19

One of the research questions of this study is whether there is a relationship between performance performance and fNIR results. Correlations show us there is no strong relationship between behavior and brain activations. In the literature, there are not any previous studies or any evidence for their relationship because it is not expected. Showing high performance in a task does not mean she/he shows high brain activation or vice versa. Thus, this result is not surprising.

4.8. Personality Traits

A Turkish culture based basic personality traits inventory (Gencoz and Oncul, 2012) was used in our study to observe personality differences. The test investigated a person under 6 different dimensions which were extraversion, conscientiousness, agreeableness, neuroticism, openness to experience and negative valence. The results of this questionnaire gave a value between 1 and 5 for each trait. For example, a person who had a point close to 1 for extraversion dimension, show less extraverted features while a point close to 5 means more extraverted characteristics.

In our study, we used the experimental data of Gencoz and Oncul (2012)'s study to have an idea about average value on personality trait of Turkish people. 454 participants were joined their study and descriptive statistics are shown in Table 44. The mean for each trait was taken as a base point and personality traits data in our study was classified lower or higher than that value. The value for extraversion personality trait is 3.47. If a participant had higher value than 3.47 for extraversion dimension, "higher" was written down for that trait; if she/he did not, "lower" was recorded. This higher/lower classification is given in Table 45. While agreeableness, neuroticism and openness to experience traits are divided into two groups homogeneously; more people with less negative valence and more conscientiousness are participated the experiment. Finding participants from wider area might have prevented the deviation.

Before starting any analyses, internal consistency of Basic Personality Traits Inventory was conducted by using Cronbach's alpha to ensure the reliability of participants' self-questionaries' personality traits. Cronbach's alphas for each dimension were high: 0.86 for extraversion, 0.89 for conscientiousness, 0.86 for agreeableness, 0.80 for neuroticism, 0.79 for openness to experience and 0.72 for negative valence. Results show that Cranbach's alphas were higher than 0.7 for each personality factor. Therefore, it can be concluded that they are reliable data to work on.

Descriptive statistics for personality traits in our study were also computed to see how participants of our study differ from ones at Gencoz and Oncul (2012). It is seen that mean values of traits are very close to each other. Descriptive statistics for personality trait in our experiment are given in Appendix P. Detailed results for each dimension and for each participant, and higher and lower version of results can also be seen in Appendix R and Appendix S, respectively.

Table 44: Descriptive Statistics for Personality Traits in Turkey (Gencoz and Oncul, 2012)

	N	Minimum	Maximum	Mean	Std. Deviation
Extraversion	454	1.25	5.00	3.47	0.82
Conscientiousness	454	1.13	5.00	3.45	0.75
Agreeableness	454	2.25	5.00	4.13	0.51
Neuroticism	454	1.00	4.78	2.78	0.74
Openness	454	1.33	5.00	3.66	0.68
Negative valence	454	1.00	4.17	1.69	0.55

Table 45: Higher and Lower Classification for Personality in the Experiment

Personality Trait	N		N
Extraversion	37	lower	16
		higher	21
Conscientiousness	37	lower	24
		higher	13
Agreeableness	37	lower	20
		higher	17
Neuroticism	37	lower	19
		higher	18
Openness	37	lower	18
		higher	19
Negative valence	37	lower	27
		higher	10

4.9. Personality Differences in Performance Measures

Pearson correlation coefficient was calculated to find out whether there is a correlation between personality traits (extraversion, conscientiousness, agreeableness, neuroticism, openness to experience and negative valence) and performance measures (maze path length, maze time, IST-1 completion time, IST-1 performance, IST-1 trial number, IST-2 completion time, IST-2 performance, IST-2 trial number and n-back performances).

First, the correlation was conducted for maze performance measures. Results show that we are not confident that there is a correlation between any personality trait and any maze performance measures. Test summary and group statistics of each personality trait for each maze performance measure can be found in Table 46 and Table 47, respectively.

When the relationship between personality traits and maze task is searched in the literature, results show that no correlation between maze scores and two personality traits (emotional stability and sociability) was found in Davies (1965). Emotional stability stands for negative meaning of neuroticism in the personality traits that was used in our experiment. Sociability has closer meaning to extraversion and openness to experience traits. Therefore, literature has similar finding with our study about their relationship.

Table 46: Test Summary of All Personality Traits for Maze Performance Measures

		Maze Path	Maze Time
PF1 (Extraversion)	Pearson Correlation	.055	.026
	Sig. (2-tailed)	.745	.879
	N	37	37
PF2 (Conscientiousness)	Pearson Correlation	-.173	-.066
	Sig. (2-tailed)	.307	.700
	N	37	37
PF3 (Agreableness)	Pearson Correlation	.020	.134
	Sig. (2-tailed)	.908	.428
	N	37	37
PF4 (Neuroticism)	Pearson Correlation	.038	.089
	Sig. (2-tailed)	.822	.602
	N	37	37
PF5 (Openness to Experience)	Pearson Correlation	-.245	-.305
	Sig. (2-tailed)	.143	.067
	N	37	37
PF6 (Negative Valence)	Pearson Correlation	-.006	-.205
	Sig. (2-tailed)	.973	.225
	N	37	37

Table 47: Group Statistics of All Personality Traits for Maze Performance Measures

Personality Trait	Measurement		N	Mean	Std. Deviation	Std. Error Mean
PF1 (Extraversion)	MazePath	lower	16	189	129.178	32.295
		higher	21	198.1	122.395	26.709
	MazeTime	lower	16	218.02	127.239	31.81
		higher	21	239.52	99.988	21.819
PF2 (Conscientiousness)	MazePath	lower	24	214.25	131.07	26.755
		higher	13	157.08	103.271	28.642
	MazeTime	lower	24	243.05	124.213	25.355
		higher	13	206.54	82.104	22.772
PF3 (Agreeableness)	MazePath	lower	20	181.33	116.978	26.157
		higher	17	209.27	133.112	32.284
	MazeTime	lower	20	207.1	83.939	18.769
		higher	17	257.43	134.595	32.644
PF4 (Neuroticism)	MazePath	lower	19	175.03	129.227	29.647
		higher	18	214.36	117.767	27.758
	MazeTime	lower	19	211.89	113.237	25.978
		higher	18	249.57	109.294	25.761
PF5 (Openness to Experience)	MazePath	lower	18	220.7	126.365	29.785
		higher	19	169.03	118.915	27.281
	MazeTime	lower	18	249.79	112.916	26.615
		higher	19	211.68	109.753	25.179
PF6 (Negative Valence)	MazePath	lower	27	205.45	130.71	25.155
		higher	10	163.7	102.058	32.274
	MazeTime	lower	27	249.49	116.049	22.334
		higher	10	178.2	81.483	25.767

Pearson correlation was run for all IST performance variables (IST-1 completion time, IST-1 performance, IST-1 trial number, IST-2 completion time, IST-2 performance, IST-2 trial number) to see whether there is a correlation between any IST performance measure and any personality trait. Since $r(37) = -.398$, $p \leq .05$; we may conclude that a negative correlation between conscientiousness and IST-1 performance exists. There is also a negative correlation between agreeableness and IST-1 trial number, where $r(37) = -.341$, $p \leq .05$.

For other IST performance measures and personality traits, we cannot say that there is a correlation between them. Test summary and group statistics of each personality trait for each IST performance measure can be found in Table 48 and Table 49, respectively.

Trial number in IST show risk-taking measure during a decision making task. According to Lauriola and Levin (2001)'s results, people with low neuroticism and people with high openness to experience took more risk in a decision making task, while the correlation coefficients were not significant between risk-taking and other personality traits (extraversion, conscientiousness, agreeableness). However, in our experiment more agreeable people do more trials than others to be sure for the result, they are risk-avoiders.

Hooper *et* al. (2008) and Byrne (2015) found that there was a negative relationship between neuroticism and decision-making task performance. In our study, neuroticism trait does not have any correlation with IST performance, but less conscientious people show better performance.

Table 48: Test Summary of All Personality Traits for IST Performance Measures

		IST 1 - Performance	IST 2 - Performance	IST1 Time	IST2 Time	IST1 TrialNo	IST2 TrialNo
PF1 (Extraversion)	Pearson Correlation	-.239	.103	.088	.013	-.082	-.202
	Sig. (2-tailed)	.154	.542	.606	.941	.631	.230
	N	37	37	37	37	37	37
PF2 (Conscientiousness)	Pearson Correlation	-,398*	.123	.228	.155	.026	-.148
	Sig. (2-tailed)	.015	.469	.174	.360	.878	.381
	N	37	37	37	37	37	37
PF3 (Agreableness)	Pearson Correlation	-.084	.148	-.286	-.190	-,341*	-.153
	Sig. (2-tailed)	.619	.383	.086	.259	.039	.364
	N	37	37	37	37	37	37
PF4 (Neuroticism)	Pearson Correlation	.122	.022	.202	.084	.156	.074
	Sig. (2-tailed)	.471	.895	.231	.622	.356	.662
	N	37	37	37	37	37	37
PF5 (Openness to Experience)	Pearson Correlation	-.027	-.042	-.141	-.098	-.088	-.142
	Sig. (2-tailed)	.873	.807	.405	.566	.603	.400
	N	37	37	37	37	37	37
PF6 (Negative Valence)	Pearson Correlation	-.081	-.063	.062	-.078	.041	-.017
	Sig. (2-tailed)	.633	.710	.717	.647	.810	.920
	N	37	37	37	37	37	37

Table 49: Group Statistics of All Personality Traits for IST Performance Measures

Personality Trait	Measurement		N	Mean	Std. Deviation	Std. Error Mean
PF1 (Extraversion)	IST1 - Performance	lower	16	1031.25	177.834	44.459
		higher	21	952.38	227.198	49.579
	IST1 - Time	lower	16	153872	66390.64	16597.7
		higher	21	165839	58657.71	12800.2
	IST1 - TrialNo	lower	16	171.75	54.773	13.693
		higher	21	166.76	54.176	11.822
	IST2 - Performance	lower	16	1081.88	328.638	82.159
		higher	21	936.19	340.786	74.365
	IST2 - Time	lower	16	162058	69141.4	17285.4
		higher	21	159882	71733.1	15653.4
	IST2 - TrialNo	lower	16	137.75	40.575	10.144
		higher	21	131.14	44.296	9.666
PF2 (Conscientiousness)	IST1 - Performance	lower	24	1016.67	194.862	39.776
		higher	13	930.77	228.709	63.432
	IST1 - Time	lower	24	149994	57797.32	11797.8
		higher	13	180363	65588.17	18190.9
	IST1 - TrialNo	lower	24	170.83	55.645	11.359
		higher	13	165.38	52.01	14.425
	IST2 - Performance	lower	24	986.67	326.132	66.571
		higher	13	1022.31	373.991	103.726
	IST2 - Time	lower	24	159979	69212.15	14127.9
		higher	13	162381	73278.99	20323.9
	IST2 - TrialNo	lower	24	139.83	41.281	8.427
		higher	13	123.23	43.603	12.093
PF3 (Agreeableness)	IST1 - Performance	lower	20	965	203.328	45.465
		higher	17	1011.76	217.607	52.778
	IST1 - Time	lower	20	172745	69119.88	15455.7
		higher	17	146452	49437.06	11990.2
	IST1 - TrialNo	lower	20	180.6	50.983	11.4
		higher	17	155.18	55.115	13.367
	IST2 - Performance	lower	20	966	355.252	79.437
		higher	17	1038.24	324.966	78.816
	IST2 - Time	lower	20	177164	77678.02	17369.3
		higher	17	141598	54997.91	13339
	IST2 - TrialNo	lower	20	136.15	45.383	10.148
		higher	17	131.47	39.529	9.587

Table 49 (continued): Group Statistics of All Personality Traits for IST Performance Measures

Personality Trait	Measurement		N	Mean	Std. Deviation	Std. Error Mean
PF4 (Neuroticism)	IST1 - Performance	lower	19	963.16	208.728	47.885
		higher	18	1011.11	211.128	49.763
	IST1 - Time	lower	19	150905	65525.56	15032.6
		higher	18	170965	56996.58	13434.2
	IST1 - TrialNo	lower	19	168.42	57.909	13.285
		higher	18	169.44	50.618	11.931
	IST2 - Performance	lower	19	968.42	366.14	83.998
		higher	18	1031.67	314.872	74.216
	IST2 - Time	lower	19	165155	79736.62	18292.8
		higher	18	156251	59154.95	13943
	IST2 - TrialNo	lower	19	133.21	46.413	10.648
		higher	18	134.83	38.756	9.135
PF5 (Openness to Experience)	IST1 - Performance	lower	18	983.33	194.785	45.911
		higher	19	989.47	225.819	51.806
	IST1 - Time	lower	18	161640	67067.69	15808
		higher	19	159740	57605.05	13215.5
	IST1 - TrialNo	lower	18	170.94	52.275	12.321
		higher	19	167	56.436	12.947
	IST2 - Performance	lower	18	1034.44	347.431	81.89
		higher	19	965.79	336.639	77.23
	IST2 - Time	lower	18	160047	67128.84	15822.4
		higher	19	161558	73799.97	16930.9
	IST2 - TrialNo	lower	18	131.61	35.119	8.278
		higher	19	136.26	48.971	11.235
PF6 (Negative Valence)	IST1 - Performance	lower	27	1018.52	184.051	35.421
		higher	10	900	253.859	80.277
	IST1 - Time	lower	27	158194	62253.86	11980.8
		higher	10	167334	62233.41	19679.9
	IST1 - TrialNo	lower	27	171.04	54.852	10.556
		higher	10	163.2	52.971	16.751
	IST2 - Performance	lower	27	989.26	348.137	66.999
		higher	10	1026	329.046	104.053
	IST2 - Time	lower	27	163940	69501.1	13375.5
		higher	10	152407	73107.72	23118.7
	IST2 - TrialNo	lower	27	136.63	39.783	7.656
		higher	10	126.9	49.992	15.809

Lastly, Pearson correlation coefficient was calculated to find out whether there is a correlation between n-back performances and personality traits. Results show that it cannot be said that there is a correlation between any personality trait and any n-back performance measures since all p values are higher than .05. Test summary and group statistics of each personality trait for each n-back performance measure can be found in Table 50 and Table 51, respectively.

Table 50: Test Summary of All Personality Traits for N-back Performance Measures

		back1	back2	back3	back4
PF1 (Extraversion)	Pearson Correlation	-.027	.094	-.082	-.119
	Sig. (2-tailed)	.875	.580	.630	.481
	N	37	37	37	37
PF2 (Conscientiousness)	Pearson Correlation	.048	.094	.133	-.293
	Sig. (2-tailed)	.778	.582	.431	.079
	N	37	37	37	37
PF3 (Agreableness)	Pearson Correlation	.294	.126	-.037	.103
	Sig. (2-tailed)	.078	.456	.829	.543
	N	37	37	37	37
PF4 (Neuroticism)	Pearson Correlation	-.063	.128	-.041	-.137
	Sig. (2-tailed)	.713	.450	.808	.418
	N	37	37	37	37
PF5 (Openness to Experience)	Pearson Correlation	.122	.122	.020	.136
	Sig. (2-tailed)	.473	.472	.904	.421
	N	37	37	37	37
PF6 (Negative Valence)	Pearson Correlation	.323	.190	-.044	.248
	Sig. (2-tailed)	.051	.261	.795	.139
	N	37	37	37	37

Table 51: Group Statistics of All Personality Traits for N-back Performance Measures

Personality Trait	Measurement		N	Mean	Std. Deviation	Std. Error Mean
PF1 (Extraversion)	1-back	lower	16	93.19	12.608	3.152
		higher	21	91.05	14.03	3.062
	2-back	lower	16	76.56	22.085	5.521
		higher	21	78.19	20.665	4.51
	3-back	lower	16	44.69	21.039	5.26
		higher	21	40.95	18.906	4.126
	4-back	lower	16	34.13	18.04	4.51
		higher	21	30.1	16.891	3.686
PF2 (Conscientiousness)	1-back	lower	24	91.13	13.911	2.839
		higher	13	93.54	12.461	3.456
	2-back	lower	24	74.63	22.604	4.614
		higher	13	82.77	17.249	4.784
	3-back	lower	24	41.38	16.5	3.368
		higher	13	44.77	25.094	6.96
	4-back	lower	24	31.83	17.643	3.601
		higher	13	31.85	17.262	4.788
PF3 (Agreeableness)	1-back	lower	20	90.65	14.091	3.151
		higher	17	93.53	12.536	3.04
	2-back	lower	20	76.15	21.495	4.806
		higher	17	79.06	20.954	5.082
	3-back	lower	20	43.7	19.618	4.387
		higher	17	41.24	20.228	4.906
	4-back	lower	20	30.55	17.913	4.006
		higher	17	33.35	16.893	4.097

Table 51 (continued): Group Statistics of All Personality Traits for N-back Performance Measures

Personality Trait	Measurement		N	Mean	Std. Deviation	Std. Error Mean
PF4 (Neuroticism)	1-back	lower	19	92	12.953	2.972
		higher	18	91.94	14.023	3.305
	2-back	lower	19	74.84	23.573	5.408
		higher	18	80.28	18.159	4.28
	3-back	lower	19	44.53	20.048	4.599
		higher	18	40.5	19.6	4.62
	4-back	lower	19	35.11	16.656	3.821
		higher	18	28.39	17.697	4.171
PF5 (Openness to Experience)	1-back	lower	18	89.61	14.508	3.42
		higher	19	94.21	11.993	2.751
	2-back	lower	18	74.94	21.74	5.124
		higher	19	79.89	20.575	4.72
	3-back	lower	18	42	20.059	4.728
		higher	19	43.11	19.81	4.545
	4-back	lower	18	29.67	18.124	4.272
		higher	19	33.89	16.643	3.818
PF6 (Negative Valence)	1-back	lower	27	90.07	14.204	2.734
		higher	10	97.1	9.171	2.9
	2-back	lower	27	76.67	23.339	4.492
		higher	10	79.7	13.549	4.284
	3-back	lower	27	44.81	18.614	3.582
		higher	10	36.5	22.117	6.994
	4-back	lower	27	30.74	14.203	2.733
		higher	10	34.8	24.426	7.724

Gray (2001) investigated extraversion and neuroticism to see whether they had any relationship with verbal and spatial 2-back tasks and found out that these traits did not have any correlations. Even though we cannot compare our study for further n-backs like 3 and 4 and with other traits (openness, conscientiousness and agreeableness), it can be said that our finding is similar with that of Gray (2001).

4.10. Personality Differences in fNIR Measures

As we mention in earlier part, there is fNIR data of 19 participants out of 37 due to physical problems. Brain activity information about left DLPFC by voxel-1 and right DLPFC by voxel-4 were used.

Pearson correlation coefficient was conducted to see whether there is a correlation between personality traits (extraversion, conscientiousness, agreeableness, neuroticism, openness to experience and negative valence) and fNIR measures (oxygenation level in voxel-1 (left DLPFC), in voxel-4 (right DLPFC)) across all 19 participants.

First, the correlation was conducted for measures in voxel-1. Results show that conscientiousness is negatively correlated with oxygenation level in voxel-1 during maze task, $r(19)= -.548$, $p\leq.05$. Extraversion ($r(19)= .459$, $p\leq.05$), agreeableness ($r(19)= .483$, $p\leq.05$), and openness to experience ($r(19)= .518$, $p\leq.05$) are all positively correlated with oxygenation level in voxel-1 during ist-2. Openness to experience also correlates positively with oxygenation level in voxel-1 during 2-back ($r(19)= .556$, $p\leq.05$) and 3-back ($r(19)= .553$, $p\leq.05$).

Correlations can be seen in Table 52. Group statistics of each personality trait (extraversion, conscientiousness, agreeableness, neuroticism, openness to experience and negative valence) for fNIR measures are given in Table 54, 55, 56, 57, 58 and 59, respectively.

Table 52: Test Summary of All Personality Traits for fNIR Measures of Left DLPFC

		V1_maze	V1_ist1	V1_ist2	V1_1back	V1_2back	V1_3back	V1_4back
PF1 (Extraversion)	Pearson Correlation	-.123	.359	,459*	.035	.206	.111	.154
	Sig. (2-tailed)	.617	.131	.048	.891	.411	.662	.541
	N	19	19	19	19	19	19	19
PF2 (Conscientiousness)	Pearson Correlation	-,548*	-.404	-.268	.034	-.117	-.188	.075
	Sig. (2-tailed)	.015	.087	.268	.895	.645	.455	.768
	N	19	19	19	19	19	19	19
PF3 (Agreableness)	Pearson Correlation	.400	.370	,483*	.203	.407	.393	.325
	Sig. (2-tailed)	.090	.119	.036	.420	.093	.106	.188
	N	19	19	19	19	19	19	19
PF4 (Neuroticism)	Pearson Correlation	-.214	-.090	-.383	.116	-.184	-.348	-.147
	Sig. (2-tailed)	.380	.715	.105	.646	.464	.157	.560
	N	19	19	19	19	19	19	19
PF5 (Openness to Experience)	Pearson Correlation	.099	.109	,518*	.103	,556*	,553*	.399
	Sig. (2-tailed)	.688	.657	.023	.685	.017	.017	.101
	N	19	19	19	19	19	19	19
PF6 (Negative Valence)	Pearson Correlation	.310	-.078	-.024	.235	.071	-.052	-.026
	Sig. (2-tailed)	.197	.751	.923	.347	.779	.838	.917
	N	19	19	19	19	19	19	19

For the relationship between oxygenation level in voxel-4 and personality traits, Pearson correlation coefficient was conducted and results show that agreeableness is positively correlated with oxygenation level in voxel-4 during ist-1, $r(19)= .560$, $p\leq .05$. Openness to experience show positive correlations with oxygenation level in voxel-4 during 2-back, $r(19)= .598$, $p\leq .05$ and 3-back, $r(19)= .609$, $p\leq .05$ and 4-back, $r(19)= .534$, $p\leq .05$.

Correlations can be seen in Table 53. Group statistics of each personality trait (extraversion, conscientiousness, agreeableness, neuroticism, openness to experience and negative valence) for fNIR measures are given in Table 54, 55, 56, 57, 58 and 59, respectively.

Table 53: Test Summary of All Personality Traits for fNIR Measures of Right DLPFC

		V4_maze	V4_ist1	V4_ist2	V4_1back	V4_2back	V4_3back	V4_4back
PF1 (Extraversion)	Pearson Correlation	.076	.406	.454	.086	.189	.101	.050
	Sig. (2-tailed)	.758	.085	.051	.735	.453	.690	.844
	N	19	19	19	18	18	18	18
PF2 (Conscientiousness)	Pearson Correlation	.002	.000	-.051	.197	.153	-.005	.125
	Sig. (2-tailed)	.994	.999	.834	.432	.543	.983	.620
	N	19	19	19	18	18	18	18
PF3 (Agreableness)	Pearson Correlation	.335	,560*	.411	.317	.265	.390	.402
	Sig. (2-tailed)	.161	.013	.081	.200	.289	.110	.098
	N	19	19	19	18	18	18	18
PF4 (Neuroticism)	Pearson Correlation	-.319	-.341	-.122	-.113	-.331	-.410	-.427
	Sig. (2-tailed)	.183	.153	.620	.656	.180	.091	.077
	N	19	19	19	18	18	18	18
PF5 (Openness to Experience)	Pearson Correlation	.160	.375	.358	.325	,598**	,609**	,534*
	Sig. (2-tailed)	.513	.113	.133	.189	.009	.007	.022
	N	19	19	19	18	18	18	18
PF6 (Negative Valence)	Pearson Correlation	-.285	-.249	-.191	.035	.081	.061	-.019
	Sig. (2-tailed)	.237	.305	.432	.890	.749	.811	.939
	N	19	19	19	18	18	18	18

Table 54: Group Statistics of Extraversion Personality Trait for fNIR Measures

Extraversion		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	lower	6	1.004	0.995	0.406
	higher	13	0.409	1.035	0.287
V4_maze	lower	6	0.343	1.029	0.42
	higher	13	0.507	0.843	0.234
V1_ist1	lower	6	-0.122	0.768	0.313
	higher	13	0.741	0.827	0.229
V4_ist1	lower	6	0.095	0.79	0.323
	higher	13	0.754	1.052	0.292
V1_ist2	lower	6	0.587	0.751	0.306
	higher	13	1.45	0.898	0.249
V4_ist2	lower	6	0.689	0.922	0.376
	higher	13	1.582	0.991	0.275
V1_1back	lower	6	0.109	1.265	0.516
	higher	13	0.05	1.064	0.295
V4_1back	lower	6	0.216	1.074	0.439
	higher	13	0.162	0.85	0.236
V1_2back	lower	6	0.802	0.789	0.322
	higher	13	1.061	1.392	0.386
V4_2back	lower	6	0.456	1.596	0.651
	higher	13	0.982	1.242	0.344
V1_3back	lower	6	1.529	1.025	0.418
	higher	13	1.563	1.515	0.42
V4_3back	lower	6	1.203	1.699	0.693
	higher	13	1.217	1.381	0.383
V1_4back	lower	6	1.701	2.059	0.841
	higher	13	1.771	1.781	0.494
V4_4back	lower	6	1.452	1.795	0.733
	higher	13	1.554	1.692	0.469

Table 55: Group Statistics of Conscientiousness Personality Trait for fNIR Measures

Conscientiousness		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	lower	13	0.74	1.074	0.298
	higher	6	0.287	0.957	0.391
V4_maze	lower	13	0.504	0.926	0.257
	higher	6	0.349	0.842	0.344
V1_ist1	lower	13	0.466	0.541	0.15
	higher	6	0.475	1.463	0.597
V4_ist1	lower	13	0.4	0.923	0.256
	higher	6	0.861	1.19	0.486
V1_ist2	lower	13	1.261	0.723	0.201
	higher	6	0.998	1.34	0.547
V4_ist2	lower	13	1.262	0.755	0.21
	higher	6	1.384	1.574	0.643
V1_1back	lower	13	-0.035	0.947	0.263
	higher	6	0.294	1.441	0.588
V4_1back	lower	13	0.041	0.739	0.205
	higher	6	0.479	1.193	0.487
V1_2back	lower	13	0.884	0.757	0.21
	higher	6	1.186	1.97	0.804
V4_2back	lower	13	0.63	1.25	0.347
	higher	6	1.22	1.561	0.637
V1_3back	lower	13	1.548	0.859	0.238
	higher	6	1.562	2.188	0.893
V4_3back	lower	13	1.098	1.148	0.318
	higher	6	1.461	2.047	0.836
V1_4back	lower	13	1.6	1.391	0.386
	higher	6	2.072	2.652	1.083
V4_4back	lower	13	1.316	1.272	0.353
	higher	6	1.968	2.423	0.989

Table 56: Group Statistics of Agreeableness Personality Trait for fNIR Measures

Agreeableness		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	lower	11	0.266	1.018	0.307
	higher	8	1.053	0.927	0.328
V4_maze	lower	11	0.083	0.678	0.205
	higher	8	0.967	0.902	0.319
V1_ist1	lower	11	0.222	0.841	0.254
	higher	8	0.809	0.889	0.314
V4_ist1	lower	11	0.2	0.836	0.252
	higher	8	1.021	1.073	0.379
V1_ist2	lower	11	1.019	1.026	0.309
	higher	8	1.396	0.789	0.279
V4_ist2	lower	11	1.182	0.981	0.296
	higher	8	1.463	1.154	0.408
V1_1back	lower	11	0.096	0.941	0.284
	higher	8	0.031	1.348	0.477
V4_1back	lower	11	0.1	0.881	0.266
	higher	8	0.288	0.966	0.342
V1_2back	lower	11	0.671	1.133	0.342
	higher	8	1.404	1.272	0.45
V4_2back	lower	11	0.629	1.591	0.48
	higher	8	1.074	0.936	0.331
V1_3back	lower	11	1.236	1.411	0.425
	higher	8	1.988	1.211	0.428
V4_3back	lower	11	0.921	1.679	0.506
	higher	8	1.613	0.993	0.351
V1_4back	lower	11	1.486	1.724	0.52
	higher	8	2.111	1.992	0.704
V4_4back	lower	11	1.174	1.785	0.538
	higher	8	2	1.488	0.526

Table 57: Group Statistics of Neuroticism Personality Trait for fNIR Measures

Neuroticism		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	lower	7	0.952	0.727	0.275
	higher	12	0.39	1.155	0.334
V4_maze	lower	7	0.812	0.814	0.308
	higher	12	0.247	0.881	0.254
V1_ist1	lower	7	0.623	1.203	0.455
	higher	12	0.379	0.692	0.2
V4_ist1	lower	7	0.732	0.839	0.317
	higher	12	0.437	1.111	0.321
V1_ist2	lower	7	1.537	1.107	0.418
	higher	12	0.968	0.784	0.226
V4_ist2	lower	7	1.366	1.216	0.46
	higher	12	1.262	0.971	0.28
V1_1back	lower	7	-0.15	1.221	0.461
	higher	12	0.196	1.05	0.303
V4_1back	lower	7	0.215	1.216	0.459
	higher	12	0.158	0.712	0.206
V1_2back	lower	7	1.108	0.926	0.35
	higher	12	0.905	1.392	0.402
V4_2back	lower	7	1.185	1.176	0.445
	higher	12	0.601	1.432	0.414
V1_3back	lower	7	1.899	0.999	0.378
	higher	12	1.35	1.522	0.439
V4_3back	lower	7	1.653	1.004	0.38
	higher	12	0.956	1.627	0.47
V1_4back	lower	7	1.823	1.892	0.715
	higher	12	1.706	1.853	0.535
V4_4back	lower	7	1.853	1.188	0.449
	higher	12	1.329	1.926	0.556

Table 58: Group Statistics of Openness to Experience Personality Trait for fNIR Measures

Openness to Experience		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	lower	8	.329	1.077	.381
	higher	11	.792	1.007	.304
V4_maze	lower	8	.119	.919	.325
	higher	11	.699	.803	.242
V1_ist1	lower	8	.318	.581	.205
	higher	11	.579	1.074	.324
V4_ist1	lower	8	.111	.952	.337
	higher	11	.862	.960	.289
V1_ist2	lower	8	.726	.883	.312
	higher	11	1.506	.852	.257
V4_ist2	lower	8	1.011	.907	.321
	higher	11	1.511	1.113	.336
V1_1back	lower	8	.104	.929	.328
	higher	11	.043	1.247	.376
V4_1back	lower	8	-.021	.570	.202
	higher	11	.325	1.079	.325
V1_2back	lower	8	.504	.932	.329
	higher	11	1.325	1.321	.398
V4_2back	lower	8	.081	1.145	.405
	higher	11	1.351	1.254	.378
V1_3back	lower	8	1.021	1.399	.495
	higher	11	1.938	1.233	.372
V4_3back	lower	8	.451	1.414	.500
	higher	11	1.767	1.237	.373
V1_4back	lower	8	1.228	1.688	.597
	higher	11	2.128	1.887	.569
V4_4back	lower	8	.784	1.497	.529
	higher	11	2.059	1.652	.498

Table 59: Group Statistics of Negative Valence Personality Trait for fNIR Measures

Negative Valence		N	Mean	Std. Deviation	Std. Error Mean
V1_maze	lower	13	.510	1.099	.305
	higher	6	.787	.942	.385
V4_maze	lower	13	.694	.886	.246
	higher	6	-.063	.656	.268
V1_ist1	lower	13	.569	.860	.239
	higher	6	.253	.991	.405
V4_ist1	lower	13	.746	1.051	.291
	higher	6	.113	.811	.331
V1_ist2	lower	13	1.239	.895	.248
	higher	6	1.045	1.075	.439
V4_ist2	lower	13	1.478	1.082	.300
	higher	6	.915	.889	.363
V1_1back	lower	13	-.045	1.231	.341
	higher	6	.316	.761	.311
V4_1back	lower	13	.042	.902	.250
	higher	6	.478	.885	.361
V1_2back	lower	13	.974	1.370	.380
	higher	6	.992	.909	.371
V4_2back	lower	13	.611	1.033	.286
	higher	6	1.262	1.888	.771
V1_3back	lower	13	1.599	1.522	.422
	higher	6	1.450	.992	.405
V4_3back	lower	13	1.147	1.398	.388
	higher	6	1.356	1.653	.675
V1_4back	lower	13	1.758	2.009	.557
	higher	6	1.730	1.473	.601
V4_4back	lower	13	1.517	1.708	.474
	higher	6	1.533	1.761	.719

In the literature, there are very few studies on personality differences by neuroimaging. Canli *et* al. (2001) investigated the Big Five personality traits which had five personality dimensions by FMRI but using only women participants. They found out that brain activation was positively correlated with extraversion and negatively correlated with neuroticism.

Our findings show extraversion has a relationship with brain activations during IST-2; conscientiousness has a relationship with brain activations during maze, agreeableness has a relationship with brain activations during IST-1 and IST-2; openness to experience has a relationship with brain activations during IST-2, 2-back, 3-back and 4-back. However neuroticism and negative valence has no relationship with blood oxygenation level during tasks. Lack of studies on that topic makes it hard to interpret and compare our findings with literature.

CHAPTER 5

CONCLUSION

5.1. Outcomes of the Study

In this study, gender and personality differences on cognitive tasks are studied. Two types of measures are used in study. One of them is called performance measure and it depends on performance of participants, and fNIR device is used as a second measure to assess human mental workload.

There are three different cognitive tasks used in the study: N-back as a working-memory task, IST as a decision making task, maze as a spatial task. N-back has four, IST has two difficulty levels and maze has one difficulty level.

No gender difference in maze performance measures exists even though literature indicates males are significantly better than females. Lack of variety in participants and the simplicity of maze may cause this disparity. We can conclude that gender is not a criterion for spatial based tasks such as a maze task, if people have similar background according to our results. Different results can be found with participants with different specialty and/or a more complex maze task.

Gender difference on trial numbers during IST-2 is expected according to literature. It is observed that females make their choices according to their instincts, whereas males analyze data and make their choice according to facts. No gender differences do not come up for other IST performance measures. Gender do not have any effects on performance or completion time of a decision making task. We see clearly that gender is not important, if decision making performance is cared. However, if the number of trials is an essential consideration, gender should be looked into.

Males perform better than females during 4-back task, while there is no gender difference during 1, 2, and 3-back tasks. Since males are better at spatial-based working memory tasks according to literature (Coluccia & Louse, 2004; Galeo & Kimura, 1993; Persson, 2013), it is an expected result. Gender difference was observed at 2 or 3-back level on previous studies, but it runs across at 4^{th} level in our study. Similar educational level of participants may prevent gender differences at 2 and 3-backs. We can use gender as a personnel selection criterion for jobs require working memory abilities but educational and analytical background should also be taken into consideration.

fNIR results show that there is no difference in oxygenation level at both left and right DLPFC regions between men and women during all cognitive tasks. In other words, gender does not have any effects on mental workload during a cognitive task. This finding can be used in employment selections. If a job involves cognitive complexity which can create mental workload, gender is not a criterion during a personnel selection to this job for people with similar educational background.

Performance at different difficulty levels are investigated and people show less accuracy at more difficult levels which is expected according to literature. Results of fNIR tests show that rise in difficulty level of the task increases changes in oxygenation level in voxels. This finding is expected and consistent with literature. When participants work harder on the task at more difficult level, increase in oxygenation level is observed as a symptom for human mental workload.

Results of our study show that for maze task, there is not a significant personality difference in performance measures for each personality trait which is a similar finding with literature. We conclude people with different personalities do not show different performance during a spatial task like a maze task. Tasks which require spatial abilities can be performed by any person with different personalities. Personality does not affect the performance of the person. Therefore, spatial ability based jobs should find another criteria than personality test for its personnel selection.

Decision making task shows that less conscientious people show better performance during IST-1 task. People with more conscientiousness have more performance concerns and their competitive structure may cause poor performance at the start of the task. Therefore, they show worse performance at the first part of the decision making task, but they put themselves together at the second part of the task. Results also show that more agreeable people do fewer trials than others. People with more agreeableness care how they seem than how they actually are. They pay relatively more importance to others' thoughts. Since experiments were done under observations, people with more agreeableness did fewer trials during IST-1 task to show the observer that they completed their tasks fast to make observers to believe that they were good at this. Only two personality traits- agreeableness and conscientiousness- have a relationship with the decision making task during our experiments. These results show us personality can affect the decision making task performance and behavior of a person. During a personnel selection for a job based on decision making tasks, personality traits can be used as a part of the job interview process.

Results for personality differences of n-back tasks are similar with maze task. There is no significant effect of personality differences in n-back performance measures. We can conclude that different personalities do not affect short-memory task performance. If a job requires a short-memory cognitive ability, the selection of a person should be based on his/ her working side rather than his/her personality factors.

Oxygenation level differences for each personality trait are searched and it is found that oxygenation levels do not differ during maze and 1-back tasks for all personality traits. Since task difficulties of 1-back and maze tasks are lower than other cognitive tasks, no oxygenation level difference is expected. Our findings show that extraversion has a positive correlation with brain activations during IST-2; agreeableness has a positive correlation with brain activations during IST-1 and IST-2. More extraverted and/or more agreeable people mind other people's thoughts about them. Thinking about thoughts of the observer during experiments may cause the increase in their mental workload therefore they may result incremental changes in the level of blood oxygenation.

Results show that conscientiousness has a negative correlation with brain activations during the maze task. In other words, people with more conscientiousness show less mental workload during the maze task. More conscientious people are more focused on their success, so they are more oriented to tasks, ignore distractions. Handling a task more consciously may provide less mental workload to people.

According to our findings, openness to experience has a positive correlation with brain activations during IST-2, 2-back, 3-back and 4-back. More open people like to try new things and they can be enthusiastic during doing those tasks. Our experiments with all different tasks and tools may seem new and different to other people. Therefore, more open people may be excited and alerted during the experiment and that can cause the increase in blood oxygenation level. As seen from results, the oxygenation level in a brain has a relationship with personality. That shows how a personality is important and distinguishing on mental workload during doing a job. Therefore, personality traits should be involved in the personnel selection to the jobs especially to mentally challenging ones.

5.2. Limitations of the Study

In this study, even though 37 participants are recruited, data of 19 participants were available to be used after filtering functions of fNIR. That creates loss of almost

%50 of the data. Therefore, experiments part of our study lasted more than expected. Increased number of participants may give chance to reach more accurate findings.

Since three types of tasks are decided to be used, task times are short. Gender differences might have been observed, if a more difficult and time-consuming maze had been used.

All participants are undergraduate or graduate engineering students from the Middle East Technical University or Hacettepe University. Such a narrow area of participant structure can affect the results. In the future, participants from different fields and different schools should be chosen in order to obtain more reliable results.

REFERENCES

Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality, and interests: evidence for overlapping traits. Psychological Bulletin, 121(2), 219.

Agor, W. H. (1986). The logic of intuitive decision making: A research-based approach for top management. Quorum Books.

Ayaz, H., Allen, S.L., Platek, S.M., & Onaral, B. (2008). Maze Suite 1.0: a complete set of tools to prepare, present, and analyze navigational and spatial cognitive neuroscience experiments. Behavior Research Methods, 40 (1), 353-359.

Ayaz, H., Izzetoglu, M., Shewokis, P., & Onaral, B. (2010, August). Sliding-window motion artifact rejection for functional near-infrared spectroscopy. In Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE (pp. 6567-6570). IEEE.

Ayaz, H., Shewokis, P. A., Bunce, S., Izzetoglu, K., Willems, B., & Onaral, B. (2012). Optical brain monitoring for operator training and mental workload assessment. Neuroimage, 59(1), 36-47.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1), 7-15.

Bell, E. C., Willson, M. C., Wilman, A. H., Dave, S., & Silverstone, P. H. (2006). Males and females differ in brain activation during cognitive tasks. Neuroimage, 30(2), 529-538.

Blokland, G. A., McMahon, K. L., Hoffman, J., Zhu, G., Meredith, M., Martin, N. G., Thompson, P.M., Zubicaray, G.I., & Wright, M. J. (2008). Quantifying the heritability of task-related brain activation and performance during the N-back working memory task: a twin fMRI study. Biological Psychology, 79(1), 70-79.

Boles, D.B. & Adair, L.P. (2001). The multiple resources questionnaire (MRQ). Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 45 (25), 1790-1794.

Bolla, K. I., Eldreth, D. A., Matochik, J. A., & Cadet, J. L. (2004). Sex-related differences in a gambling task and its neurological correlates. Cerebral Cortex, 14(11), 1226-1232.

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neuroscience & Biobehavioral Reviews, 44, 58-75.

Borghini, G., Isabella, R., Vecchiato, G., Toppi, J., Astolfi, L., Caltagirone, C., & Babiloni, F. (2011). Brainshield: HREEG study of perceived pilot mental workload. Italian Journal of Aerospace Medicine, 5, 34-47

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. Cerebral Cortex, 11(9), 825-836.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. Biological Psychology, 42(3), 361-377.

Bunce, S. C., Izzetoglu, M., Izzetoglu, K., Onaral, B., & Pourrezaei, K. (2006). Functional near-infrared spectroscopy. Engineering in Medicine and Biology

Magazine, IEEE, 25(4), 54-62.

Byrne, K. A., Silasi-Mansat, C. D., & Worthy, D. A. (2015). Who chokes under pressure? The Big Five personality traits and decision-making under pressure. Personality and Individual Differences, 74, 22-28.

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. Behavioral Neuroscience, 115(1), 33.

Cansiz, Y., 2012. Effects of way finding affordances on usability of virtual world environments in terms of users' satisfaction, performance, and mental workload: examination by eye-tracking and fNIR device. MS. Thesis, Middle East Technical University, Ankara, Turkey.

Carlson, S., Martinkauppi, S., Rämä, P., Salli, E., Korvenoja, A., & Aronen, H. J. (1998). Distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. Cerebral Cortex, 8(8), 743-752.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. Journal of Experimental Child Psychology, 33(3), 386-404.

Chamorro-Premuzic, T., Furnham, A., & Petrides, K. (2006). Personality and intelligence. Journal of Individual Differences, 27(3), 147-150.

Chance, B., Anday, E., Nioka, S., Zhou, S., Hong, L., Worden, K., Li, C., Murray, T., Ovetsky, Y., Pidikiti, D. & Thomas, R. (1998). A novel method for fast imaging of brain function non-invasively with light. Optics Express, 2(10), 411-423.

Clark, L., Robbins, T. W., Ersche, K. D., & Sahakian, B. J. (2006). Reflection impulsivity in current and former substance users. Biological Psychiatry, 60(5), 515-522.

Colle, H. A., & Reid, G. B. (1999). Double trade-off curves with different cognitive processing combinations: Testing the cancellation axiom of mental workload measure theory. Human Factors: The Journal of the Human Factors and Ergonomics Society, 41(1), 35-50.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. Journal of Environmental Psychology, 24(3), 329-340.

Costa, P. T., Fozard, J. L., McCrae, R. R., & Bosse, R. (1976). Relations of age and personality dimensions to cognitive ability factors. Journal of Gerontology, 31(6), 663-669.

Coyle, S. M., Ward, T. E., & Markham, C. M. (2007). Brain–computer interface using a simplified functional near-infrared spectroscopy system. Journal of Neural Engineering, 4(3), 219-226.

Crone, E. A., Vendel, I., & van der Molen, M. W. (2003). Decision-making in disinhibited adolescents and adults: insensitivity to future consequences or driven by immediate reward?. Personality and Individual Differences, 35(7), 1625-1641.

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. Neuroimage, 54(4), 2808-2821.

Damos, D. L., & Bloem, K. A. (1985). Type A behavior pattern, multiple-task performance, and subjective estimation of mental workload. Bulletin of the Psychonomic Society, 23(1), 53-56.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19(4), 450-466.

Davies, A. D. (1965). The perceptual maze test in a normal population. Perceptual and Motor Skills, 20(1), 287-293.

Deco, G., & Rolls, E. T. (2005). Attention, short-term memory, and action selection: a unifying theory. Progress in Neurobiology, 76(4), 236-256.

DeYoung, C. G. (2010). Personality neuroscience and the biology of traits. Social and Personality Psychology Compass, 4(12), 1165-1180.

De Rivecourt, M., Kuperus, M. N., Post, W. J., & Mulder, L. J. M. (2008). Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. Ergonomics, 51(9), 1295-1319.

De Waard, D. (1996). The measure of drivers' mental workload. PhD thesis, University of Groningen, The Netherlands.

Dussault, C., Jouanin, J. C., Philippe, M., & Guezennec, C. Y. (2005). EEG and ECG changes during simulator operation reflect mental workload and vigilance. Aviation, Space, and Environmental Medicine, 76(4), 344-351.

Eggemeier, F.T. (1988). Properties of workload assessment techniques. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. North-Holland, Amsterdam, 41-62.

Field, A. (2009). Discovering statistics using SPSS. Sage publications.

Friedman, M., & Rosenman, R. H. (1974). Type A behavior and your heart. New York: Knopf.

Fujimoto, T., Matsumoto, T., Fujita, S., Takeuchi, K., Nakamura, K., Mitsuyama, Y., & Kato, N. (2008). Changes in glucose metabolism due to aging and gender-related differences in the healthy human brain. Psychiatry Research: Neuroimaging, 164(1), 58-72.

Furnham, A., & Chamorro-Premuzic, T. (2006). Personality, intelligence and general knowledge. Learning and Individual Differences, 16(1), 79-90.

Fuster, J. M. (2008). The prefrontal cortex. 4. China: Elsevier.

Galea, L. A., & Kimura, D. (1993). Sex differences in route-learning. Personality and Individual Differences, 14(1), 53-65.

Galley, N. (1993). The evaluation of the electrooculogram as a psychophysiological measuring instrument in the driver study of driver behaviour. Ergonomics, 36(9), 1063-1070.

Gencoz, T., & Oncul, O. (2012). Examination of Personality Characteristics in a Turkish Sample: Development of Basic Personality Traits Inventory. The Journal of General Psychology, 139(3), 194-216.

Gevins, A., Smith, M. E., Le, J., Leong, H., Bennett, J., Martin, N., McEvoy, L., Du, R. & Whitfield, S. (1996). High resolution evoked potential imaging of the cortical dynamics of human working memory. Electroencephalography and Clinical Neurophysiology, 98(4), 327-348.

Gopher, D. & Braune, R. (1984). On the psychophysics of workload-why bother with subjective measures. Human Factors, 26(5), 519–532.

Graham, E. K., & Lachman, M. E. (2012). Personality stability is associated with better cognitive performance in adulthood: Are the stable more able?. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 67(5), 545-554.

Graham, E. K., & Lachman, M. E. (2014). Personality traits, facets and cognitive performance: Age differences in their relations. Personality and Individual Differences, 59, 89-95.

Gray, J. R. (2001). Emotional modulation of cognitive control: Approach–withdrawal states double-dissociate spatial from verbal two-back task performance. Journal of Experimental Psychology: General, 130(3), 436.

Greely, H. T., & Wagner, A. D. (2011). Reference guide on neuroscience: Third Edition. Federal judicial center reference manual on scientific evidence, Federal Judicial Center, Washington, DC, 747-812.

Gron, G., Wunderlich, A. P., Spitzer, M., Tomczak, R., & Riepe, M. W. (2000). Brain activation during human navigation: gender-different neural networks as substrate of performance. Nature Neuroscience, 3(4), 404-408.

Gronwall, D. M. A. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. Perceptual and Motor Skills, 44(2), 367-373.

Hankins, T. C., & Wilson, G. F. (1998). A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. Aviation Space and Environmental Medicine, 69(4), 360-367.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in Psychology, 52, 139-183.

Harvey, P. O., Fossati, P., Pochon, J. B., Levy, R., LeBastard, G., Lehericy, S., Allilaire, J.F., & Dubois, B. (2005). Cognitive control and brain resources in major depression: an fMRI study using the n-back task. Neuroimage, 26(3), 860-869.

Haut, K. M., & Barch, D. M. (2006). Sex influences on material-sensitive functional lateralization in working and episodic memory: men and women are not all that different. Neuroimage, 32(1), 411-422.

Hooper, C. J., Luciana, M., Wahlstrom, D., Conklin, H. M., & Yarger, R. S. (2008). Personality correlates of Iowa Gambling Task performance in healthy adolescents. Personality and Individual Differences, 44(3), 598-609.

Izzetoglu, K., Ayaz, H., Merzagora, A., Izzetoglu, M., Shewokis, P. A., Bunce, S. C., Pourrezaei, K., Rosen, A. & Onaral, B. (2011). The evolution of field deployable fNIR spectroscopy from bench to clinical settings. Journal of Innovative Optical Health Sciences, 4(03), 239-250.

Izzetoglu, K., Bunce, S., Onaral, B., Pourrezaei, K., & Chance, B. (2004). Functional optical brain imaging using near-infrared during cognitive tasks. International Journal of Human-Computer Interaction, 17(2), 211-227.

Jex, H.R. (1988). Measuring mental workload: Problems, progress, and promises In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. North-Holland, Amsterdam, 5-39.

Johnson, P. (1992). Human computer interaction: psychology, task analysis, and software engineering. McGraw-Hill.

Jorna, P. G. A. M. (1993). Heart-rate and workload variations in actual and simulated flight. Ergonomics, 36(9), 1043-1054.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. Personnel Psychology, 52(3), 621-652.

Koch, K., Pauly, K., Kellermann, T., Seiferth, N. Y., Reske, M., Backes, V., Stocker, T., Shah, J.N., Amunts, K., Kircher, T., Schneider, F., & Habel, U. (2007). Gender differences in the cognitive control of emotion: An fMRI study. Neuropsychologia, 45(12), 2744-2754

Lee, D. H., Parks, K.S. (1990). Multivariate analysis of mental and physical load components in sinus arrhythmia scores. Ergonomics, 33(1), 35-47.

León-Carrión, J., Izzetoglu, M., Izzetoglu, K., Martín-Rodríguez, J. F., Damas-López, J., Martin, J. M. B., & Domínguez-Morales, M. R. (2010). Efficient learning produces spontaneous neural repetition suppression in prefrontal cortex. Behavioural Brain Research, 208(2), 502-508.

Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: An exploratory study. Personality and Individual Differences, 31(2), 215-226.

Li, T., Luo, Q., & Gong, H. (2010). Gender-specific hemodynamics in prefrontal cortex during a verbal working memory task by near-infrared spectroscopy. Behavioural Brain Research, 209(1), 148-153.

Lighthall, N. R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E. Y., Samii, N. & Mather, M. (2012). Gender differences in reward-related decision processing under stress. Social Cognitive and Affective Neuroscience, 7(4), 476-484.

Mandrick, K., Derosiere, G., Dray, G., Coulon, D., Micallef, J. P., & Perrey, S. (2013). Prefrontal cortex activity during motor tasks with additional mental load requiring attentional demand: a near-infrared spectroscopy study. Neuroscience Research, 76(3), 156-162.

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. Brain, 125(3), 624-639.

Manoach, D. S., Schlaug, G., Siewert, B., Darby, D. G., Bly, B. M., Benfield, A., Edelman, R.R. & Warach, S. (1997). Prefrontal cortex fMRI signal changes are correlated with working memory load. Neuroreport, 8(2), 545-549.

Marumo, K., Takizawa, R., Kawakubo, Y., Onitsuka, T., & Kasai, K. (2009). Gender difference in right lateral prefrontal hemodynamic response while viewing fearful faces: a multi-channel near-infrared spectroscopy study. Neuroscience Research, 63(2), 89-94.

McCrae, R. R., & Costa Jr, P. T. (1997). Personality trait structure as a human universal. American Psychologist, 52(5), 509.

Meshkati, N. & Lowewinthal, A. (1988). An eclectic and critical review of four primary mental workload assessment methods: A guide for developing a comprehensive model. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. North-Holland, Amsterdam, 251-267.

Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a “virtual” maze: Sex differences and correlation with psychometric measures of spatial ability in humans. Evolution and Human Behavior, 19(2), 73-87.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: Distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. The Journal of Neuroscience, 21, 7733– 7741.

Montello, D. R., & Pick, H. L. (1993). Integrating knowledge of vertically aligned large-scale spaces. Environment and Behavior, 25(3), 457-484.

Moutafi, J., Furnham, A., & Crump, J. (2006). What facets of openness and conscientiousness predict fluid intelligence score?. Learning and Individual Differences, 16(1), 31-42.

Mulder, G. (1979). Mental load, mental effort and attention. In: Moray, N. Mental Workload. New York, the USA, 299-325.

North, R.A., Riley, V.A. (1989). W/INDEX: A predictive model of operator workload. Applications of Human Performance Models to System Design, 81-89.

Parikh, J., Neubauer, F. F., & Lank, A. G. (1994). Intuition: The new frontier of management. Blackwell.

Parmenter, B. A., Shucard, J. L., Benedict, R. H., & Shucard, D. W. (2006).

Working memory deficits in multiple sclerosis: Comparison between the n-back task and the Paced Auditory Serial Addition Test. Journal of the International Neuropsychological Society, 12(05), 677-687.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. Nature Neuroscience, 4(1), 95-102.

O'Donnell, R.D. & Eggemeier, F.T. (1986). Workload assessment methodology. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.), Handbook of Perception and Human Performance. Oxford, England, 1-49.

O'Laughlin, E. M., & Brubaker, B. S. (1998). Use of landmarks in cognitive mapping:gen der differences in self report versus performance. Personality and Individual Differences, 24(5), 595–601.

Overman, W., Graham, L., Redmond, A., Eubank, R., Boettcher, L., Samplawski, O., & Walsh, K. (2006). Contemplation of moral dilemmas eliminates sex differences on the Iowa gambling task. Behavioral Neuroscience,120(4), 817.

Ozcan, O.O., 2012. Exploring the effects of working memory capacity, attention, and expertise on situation awareness in a flight simulation environment. MS. Thesis, Middle East Technical University, Ankara, Turkey.

Persson, J., Herlitz, A., Engman, J., Morell, A., Sjölie, D., Wikström, J., & Söderlund, H. (2013). Remembering our origin: gender differences in spatial memory are reflected in gender differences in hippocampal lateralization. Behavioural Brain Research, 256, 219-228.

Ragland, J.D., Turetsky, B.I., Gur, R.C., Gunning-Dixon, F., Turner, T., Schroeder, L., Chan, R., & Gur, R.E. (2002).Working memory for complex figures: An fMRI comparison of letter and fractal n-back tasks. Neuropsychology, 16, 370–379.

Reavis, R., & Overman, W. H. (2001). Adult sex differences on a decision-making task previously shown to depend on the orbital prefrontal cortex. Behavioral Neuroscience, 115(1), 196.

Reid, G.B. & Nygren, T.E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. Advances in Psychology, 52, 185-218.

Roscoe, A. H. (1992). Assessing pilot workload - why measure heart-rate, hrv and respiration. Biological Psychology, 34(2-3), 259-287.

Roscoe, A. H. (1993). Heart-rate as a psychophysiological measure for in-flight workload assessment. Ergonomics, 36(9), 1055-1062.

Rose, C. L., Murphy, L. B., Byard, L., & Nikzad, K. (2002). The role of the Big Five personality factors in vigilance performance and workload. European Journal of personality, 16(3), 185-200.

Sadalla, E. K., & Montello, D. R. (1989). Remembering changes in direction. Environment and Behavior, 21(3), 346-363.

Schmidt, H., Jogia, J., Fast, K., Christodoulou, T., Haldane, M., Kumari, V., & Frangou, S. (2009). No gender differences in brain activation during the N-back task: An fMRI study in healthy individuals. Human Brain Mapping, 30(11), 3609-3615.

Simon, H. A. (1977). The logic of heuristic decision making. In Models of Discovery (pp. 154-175). Springer Netherlands.

Smith, E. E., & Kosslyn, S. M. (2009). Cognitive Psychology. Mind and Brain.

Sohn, S. Y., & Jo, Y. K. (2003). A study on the student pilot's mental workload due to personality types of both instructor and student. Ergonomics, 46(15), 1566-1577.

Solowij, N., Jones, K. A., Rozman, M. E., Davis, S. M., Ciarrochi, J., Heaven, P. C., Pesa, N., Lubman, D.I., & Yücel, M. (2012). Reflection impulsivity in adolescent cannabis users: a comparison with alcohol-using and non-substance-using adolescents. Psychopharmacology, 219(2), 575-586.

Speck, O., Ernst, T., Braun, J., Koch, C., Miller, E., & Chang, L. (2000). Gender differences in the functional organization of the brain for working memory. Neuroreport, 11(11), 2581-2585.

Stoltenberg, S. F., & Vandever, J. M. (2010). Gender moderates the association between 5-HTTLPR and decision-making under ambiguity but not under risk. Neuropharmacology, 58(2), 423-428.

Taylor, H. A., & Tversky, B. (1992). Descriptions and depictions of environments. Memory & Cognition, 20(5), 483-496.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent?. Journal of Memory and Language, 28(2), 127-154.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. Behavior Research Methods, 37(3), 498-505.

van den Bos, R., Homberg, J., & de Visser, L. (2013). A critical review of sex differences in decision-making tasks: focus on the Iowa Gambling Task. Behavioural Brain Research, 238, 95-108.

Vicente, K. J., Thornton, D. C., & Moray, N. (1987). Spectral analysis of sinus arrhythmia: A measure of mental effort. Human factors: The Journal of the Human Factors and Ergonomics Society, 29(2), 171-182.

Vidulich, M.A. & Tsang, P.S. (1987). Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 31(9), 1057-1061.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. Psychological Bulletin, 117(2), 250.

Waller, D., Knapp, D., & Hunt, E. (2001). Spatial representations of virtual mazes: The role of visual fidelity and individual differences. Human Factors: The Journal of the Human Factors and Ergonomics Society, 43(1), 147-158.

Wang, Y., & Ruhe, G. (2007). The cognitive process of decision making.

Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. Psychophysiology, 38(06), 998-1003.

Wijers, A. A., Lange, J. J., Mulder, G., & Mulder, L. J. (1997). An ERP study of visual spatial attention and letter target detection for isoluminant and nonisoluminant stimuli. Psychophysiology, 34(5), 553-565.

Wierwille, W.W. & Casali, J.G. (1983). A validated rating scale for global mental workload measure applications. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 27(2), 129-133.

Willis, M. W., Ketter, T. A., Kimbrell, T. A., George, M. S., Herscovitch, P., Danielson, A. L., Benson, B.E. & Post, R. M. (2002). Age, sex and laterality effects on cerebral glucose metabolism in healthy adults. Psychiatry Research: Neuroimaging, 114(1), 23-37.

Wilson, G. F., Fullenkamp, P., & Davis, I. (1994). Evoked-potential, cardiac, blink, and respiration measures of pilot workload in air-to-ground missions. Aviation Space and Environmental Medicine, 65(2), 100-10

Wilson, G. F., & O'Donnell, R. D. (1988). Measure of operator workload with the neuropsychological workload test battery. Advances in Psychology, 52, 63-100.

Wilson, G. F., & Russell, C. A. (2004). Psychophysiologically determined adaptive aiding in a simulated UCAV task. Human performance, situation awareness, and automation: Current Research and Trends, 200-204.

Yang, H., Zhou, Z., Liu, Y., Ruan, Z., Gong, H., Luo, Q., & Lu, Z. (2007). Gender difference in hemodynamic responses of prefrontal area to emotional stress by near-infrared spectroscopy. Behavioural Brain Research, 178(1), 172-176.

APPENDIX A

APPROVAL BY THE HUMAN SUBJECT ETHICS COMMITTEE

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

Sayı: 28620816/ 276 - 1057

29.05.2014

Gönderilen : Prof. Dr. Canan Çilingir
Endüstri Mühendisliği Bölümü

Gönderen : Prof. Dr. Canan Özgen
IAK Başkanı

İlgi : Etik Onayı

Danışmanlığını yapmış olduğunuz Endüstri Mühendisliği Bölümü öğrencisi Betül Batun'un "Karar Verme ve Kısa Dönemli Hafıza Gerektiren Görevlerde Beyin Aktivitesinin İncelenmesi" isimli araştırması "İnsan Araştırmaları Komitesi" tarafından uygun görülerek gerekli onay verilmiştir.

Bilgilerinize saygılarımla sunarım.

Etik Komite Onayı

Uygundur

29/05/2014

Prof.Dr. Canan Özgen
Uygulamalı Etik Araştırma Merkezi
(UEAM) Başkanı
ODTÜ 06531 ANKARA

APPENDIX B

PARTICIPANTS' PERSONAL DATA

Participant No	Gender	School	Age
1	Female	Hacettepe	23
2	Male	Hacettepe	24
3	Male	Hacettepe	21
4	Male	Hacettepe	26
5	Male	Hacettepe	26
6	Female	Hacettepe	26
7	Male	Hacettepe	26
8	Female	METU	24
9	Male	METU	24
10	Female	METU	24
11	Male	METU	25
12	Female	METU	31
13	Male	METU	31
14	Female	METU	26
15	Male	METU	26
16	Female	Hacettepe	25
17	Male	METU	26
18	Female	Hacettepe	25
19	Male	METU	25
20	Male	METU	25
21	Female	METU	24
22	Female	Hacettepe	24
23	Female	METU	23
24	Female	METU	31
25	Male	METU	28
26	Male	Hacettepe	26
27	Female	METU	23
28	Female	METU	26
29	Female	METU	23
30	Female	METU	24
31	Female	Hacettepe	25
32	Female	Hacettepe	26
33	Female	Hacettepe	26
34	Female	Hacettepe	26
35	Female	METU	31
36	Female	METU	24
37	Female	METU	25

APPENDIX C

PERSONALITY QUESTIONNAIRE

Türk Kültüründe Geliştirilmiş Temel Kişilik Özellikleri Ölçeği

Yönerge: Aşağıda size uyan ya da uymayan pek çok kişilik özelliği bulunmaktadır. Bu özelliklerden her birinin sizin için ne kadar uygun olduğunu ilgili rakamı daire içine alarak belirtiniz: (1) Hiç uygun değil, (2) Uygun değil, (3) Kararsızım, (4) Uygun, (5) Çok uygun.

1	Aceleci	1	2	3	4	5	24	Pasif	1	2	3	4	5
2	Yapmacık	1	2	3	4	5	25	Disiplinli	1	2	3	4	5
3	Duyarlı	1	2	3	4	5	26	Açgözlü	1	2	3	4	5
4	Konuşkan	1	2	3	4	5	27	Sinirli	1	2	3	4	5
5	Kendine güvenen	1	2	3	4	5	28	Canayakın	1	2	3	4	5
6	Soğuk	1	2	3	4	5	29	Kızgın	1	2	3	4	5
7	Utangaç	1	2	3	4	5	30	Sabit fikirli	1	2	3	4	5
8	Paylaşımcı	1	2	3	4	5	31	Görgüsüz	1	2	3	4	5
9	Geniş / rahat	1	2	3	4	5	32	Durgun	1	2	3	4	5
10	Cesur	1	2	3	4	5	33	Kaygılı	1	2	3	4	5
11	Agresif(Saldırgan)	1	2	3	4	5	34	Terbiyesiz	1	2	3	4	5
12	Çalışkan	1	2	3	4	5	35	Sabırsız	1	2	3	4	5
13	İçten pazarlıklı	1	2	3	4	5	36	Yaratıcı	1	2	3	4	5
14	Girişken	1	2	3	4	5	37	Kaprisli	1	2	3	4	5
15	İyi niyetli	1	2	3	4	5	38	İçine kapanık	1	2	3	4	5
16	İçten	1	2	3	4	5	39	Çekingen	1	2	3	4	5
17	Kendinden emin	1	2	3	4	5	40	Alıngan	1	2	3	4	5
18	Huysuz	1	2	3	4	5	41	Hoşgörülü	1	2	3	4	5
19	Yardımsever	1	2	3	4	5	42	Düzenli	1	2	3	4	5
20	Kabiliyetli	1	2	3	4	5	43	Titiz	1	2	3	4	5
21	Üşengeç	1	2	3	4	5	44	Tedbirli	1	2	3	4	5
22	Sorumsuz	1	2	3	4	5	45	Azimli	1	2	3	4	5
23	Sevecen	1	2	3	4	5							

APPENDIX D

ENGLISH VERSION OF THE PERSONALITY QUESTIONNAIRE

Turkish Culture Based Basic Personality Traits Inventory

Instructions: There are many personality items below that fit or not fit you. Circle the number for each item on how much you agree that it is appropriate for you on a five point scale: (1) very inaccurate (2) inaccurate (3) unsure (4) accurate and (5) very accurate.

1	Impetuous	1	2	3	4	5	24	Passive	1	2	3	4	5
2	Pretentious	1	2	3	4	5	25	Self-disciplined	1	2	3	4	5
3	Sensitive	1	2	3	4	5	26	Greedy	1	2	3	4	5
4	Talkative	1	2	3	4	5	27	Nervous	1	2	3	4	5
5	Self-assured	1	2	3	4	5	28	Genial	1	2	3	4	5
6	Cold	1	2	3	4	5	29	Angry	1	2	3	4	5
7	Shy	1	2	3	4	5	30	Hidebound	1	2	3	4	5
8	Sharer	1	2	3	4	5	31	Ill-mannered	1	2	3	4	5
9	Easy-going	1	2	3	4	5	32	Lethargic	1	2	3	4	5
10	Brave	1	2	3	4	5	33	Worried	1	2	3	4	5
11	Aggressive	1	2	3	4	5	34	Rude	1	2	3	4	5
12	Hard-working	1	2	3	4	5	35	Impatient	1	2	3	4	5
13	Backstabbing	1	2	3	4	5	36	Creative	1	2	3	4	5
14	Enterprising	1	2	3	4	5	37	Capricious	1	2	3	4	5
15	Well intentioned	1	2	3	4	5	38	Withdrawn	1	2	3	4	5
16	Sincere	1	2	3	4	5	39	Timid	1	2	3	4	5
17	Self-confident	1	2	3	4	5	40	Touchy	1	2	3	4	5
18	Temperamental	1	2	3	4	5	41	Tolerant	1	2	3	4	5
19	Philanthropic	1	2	3	4	5	42	Tidy	1	2	3	4	5
20	Capable	1	2	3	4	5	43	Fussy	1	2	3	4	5
21	Lazy	1	2	3	4	5	44	Prudent	1	2	3	4	5
22	Irresponsible	1	2	3	4	5	45	Determined	1	2	3	4	5
23	Compassionate	1	2	3	4	5							

APPENDIX E

THE EVALUATION OF BASIC PERSONALITY TRAITS INVENTORY

Extraversion	Conscientiousness	Agreeableness	Neuroticism	Openness To Experience	Negative Valence
4	12	3	1	5	2
R 6	R 21	8	11	9	13
R 7	R 22	15	18	10	26
14	25	16	27	17	30
R 24	42	19	29	20	31
R 32	43	23	33	36	34
R 38	44	28	35		
R 39	45	41	37		
			40		

R: Reversing Entry

APPENDIX F

INFORMED CONSENT FORM

Gönüllü Katılım Formu

Bu çalışma, Betul BATUN, Prof. Dr. Canan ÇİLİNGİR tarafından farklı bilişsel görevlerde cinsiyet farkını çıkarmaya yönelik bir çalışmadır. Bu kapsamda üniversite öğrencilerinden bilgi toplanılması hedeflenmektedir.

Çalışmaya katılım tamamen gönüllülük temelinde olmalıdır. Deney öncesi, sizden kimlik belirleyici hiçbir bilgi istenmemektedir. Cevaplarınız tamamen gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir; elde edilecek bilgiler bilimsel yayınlarda kullanılacaktır.

Deney suresince zihinsel yorgunluğu ölçmek için fonksiyonel yakın kızılötesi spektroskopi(fNIR) adi verilen beyin görüntüleme sistemi kullanilacaktir. Beyin aktivitesi sirasindaki hemodinamik degisimleri olcen arac, kisinin alnini kaplayan esnek sensor ve sensor uzerinde bulunan 4 isik kaynagi, 10 dedektor ve 16 vokselden olusur. (bknz. Sekil-A). Bu cihaz, deney sirasinda veya sonrasinda insan sagligini tehdit edecek herhangi bir unsur icermemektedir.

Sekil-A

Deney, genel olarak kişisel rahatsızlık verecek sorular, aktiviteler içermemektedir. Ancak, katılım sırasında sorulardan ya da herhangi başka bir nedenden ötürü kendinizi rahatsız hissederseniz cevaplama işini veya deneyi yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda deneyi uygulayan kişiye, deneyi tamamlamadığınızı söylemeniz yeterli olacaktır. Deney sonunda, bu çalışmayla ilgili sorularınız cevaplandırılacaktır. Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için ODTÜ Endüstri Mühendisliği öğretim üyelerinden Prof. Dr. Canan ÇİLİNGİR (Tel: +90 (312) 210-2272 E-posta: cilingir@metu.edu.tr) ya da Betul BATUN (Tel: +90 539 252 78 51 E-posta: betulbatun@gmail.com) ile iletişim kurabilirsiniz.

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda kullanılmasını kabul ediyorum. (Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

İsim Soyad	Tarih	İmza	Alınan Ders
	----/----/-----		

APPENDIX G

MAZE COMPLETION TIME

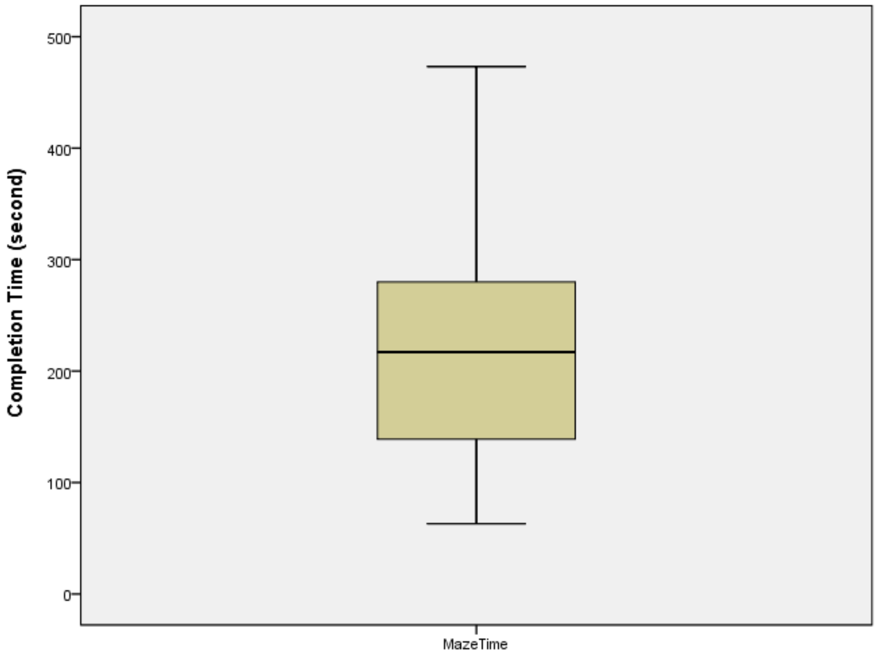

Figure 23: Boxplot of Maze Completion Time

APPENDIX H

MAZE PATH LENGTH

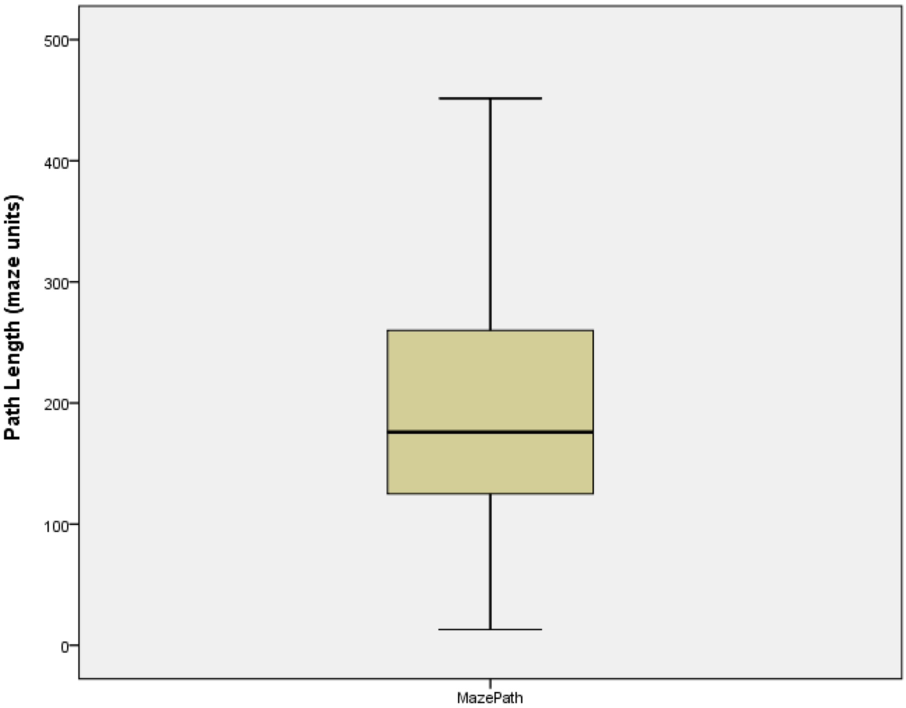

Figure 24: Boxplot of Maze Path Length

APPENDIX I

IST PERFORMANCE

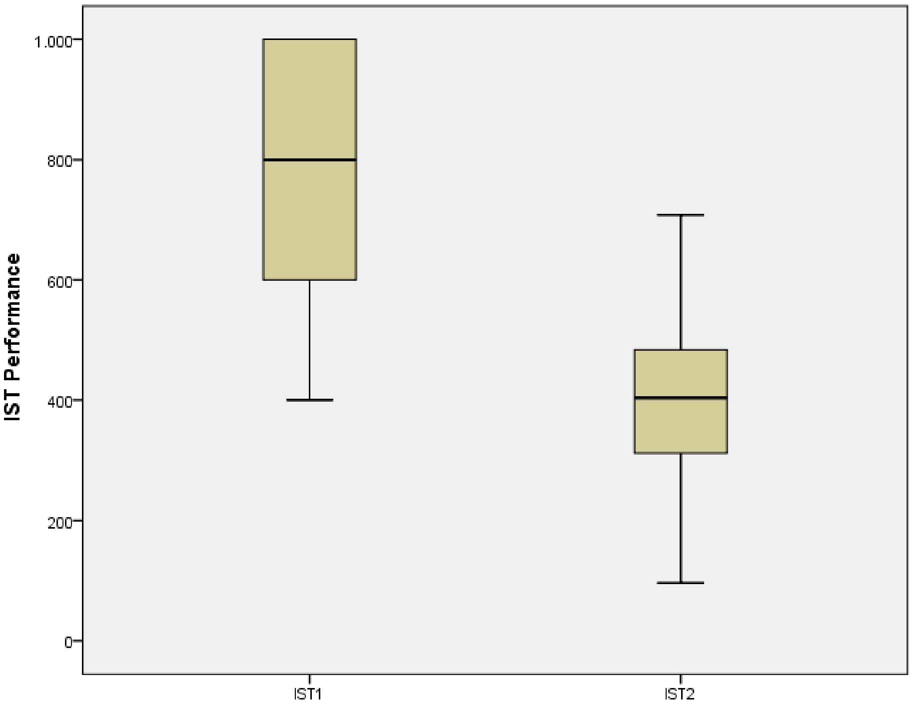

Figure 25: Boxplot of IST Performance

APPENDIX J

IST COMPLETION TIME

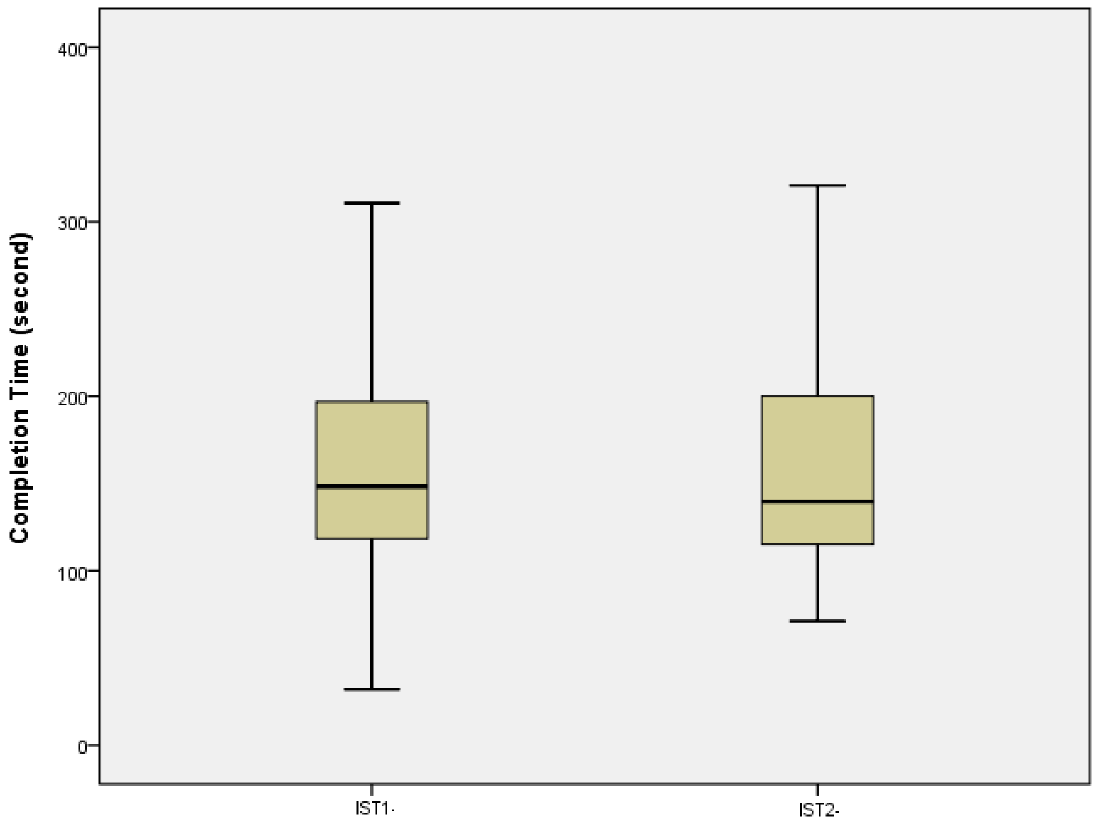

Figure 26: Boxplot of IST Completion Time

APPENDIX K

IST TRIAL NUMBERS

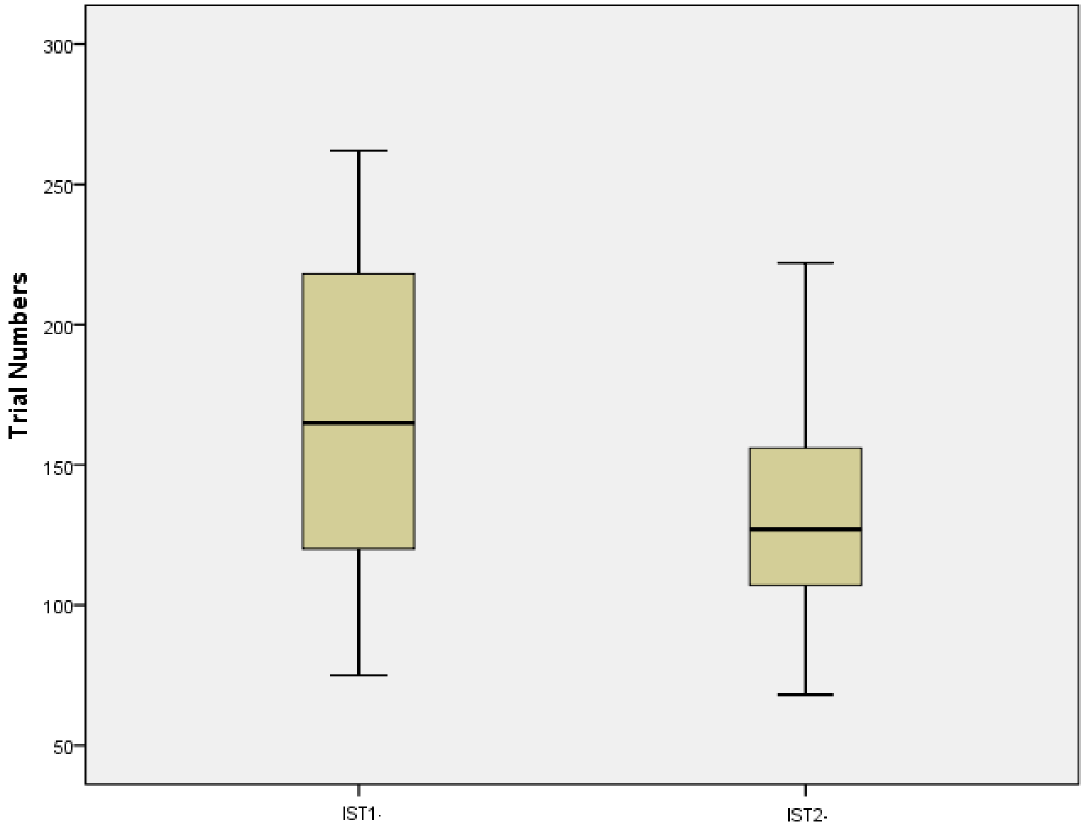

Figure 27: Boxplot of IST Trial Numbers

APPENDIX L

N-BACK PERFORMANCES

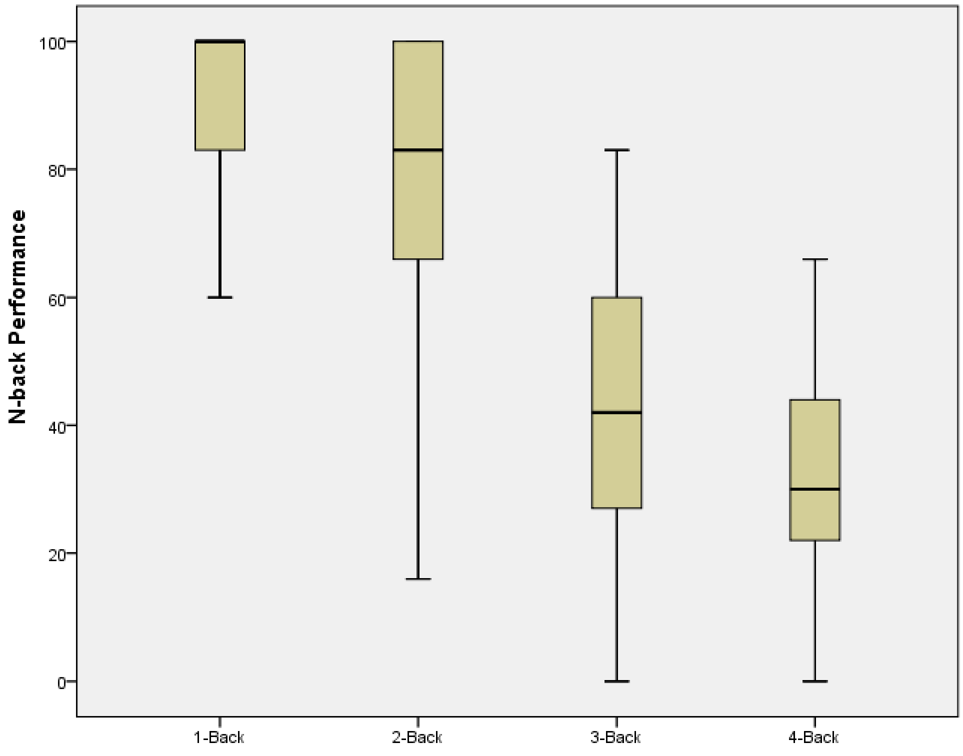

Figure 28: Boxplot of N-back Performances

APPENDIX M

FNIR MEASURES DURING MAZE TASK

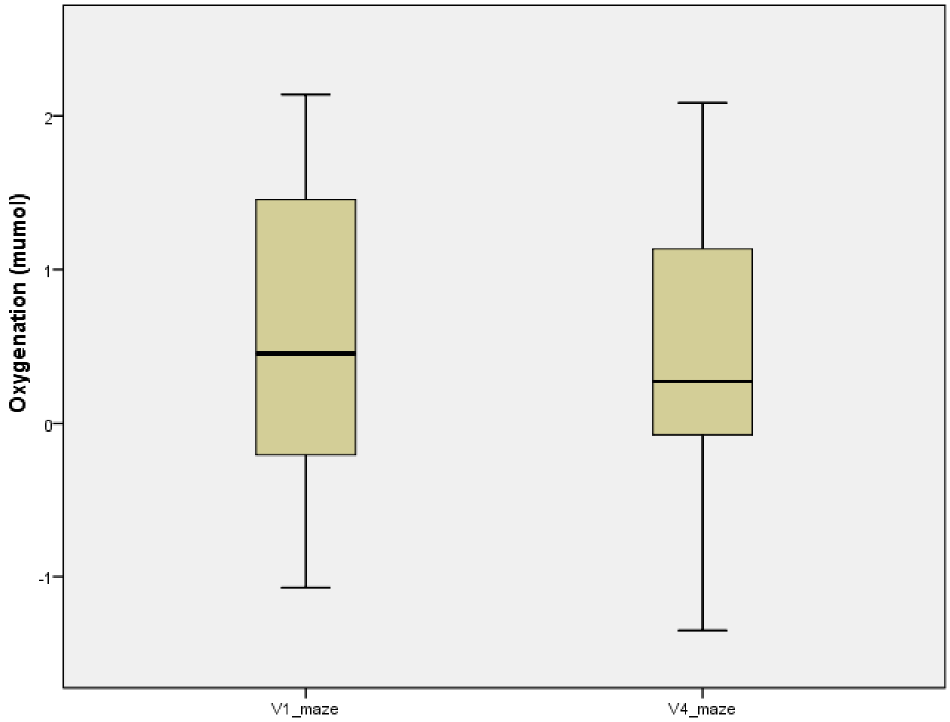

Figure 29: Boxplot of Maze fNIR Measures

APPENDIX N

FNIR MEASURES DURING IST

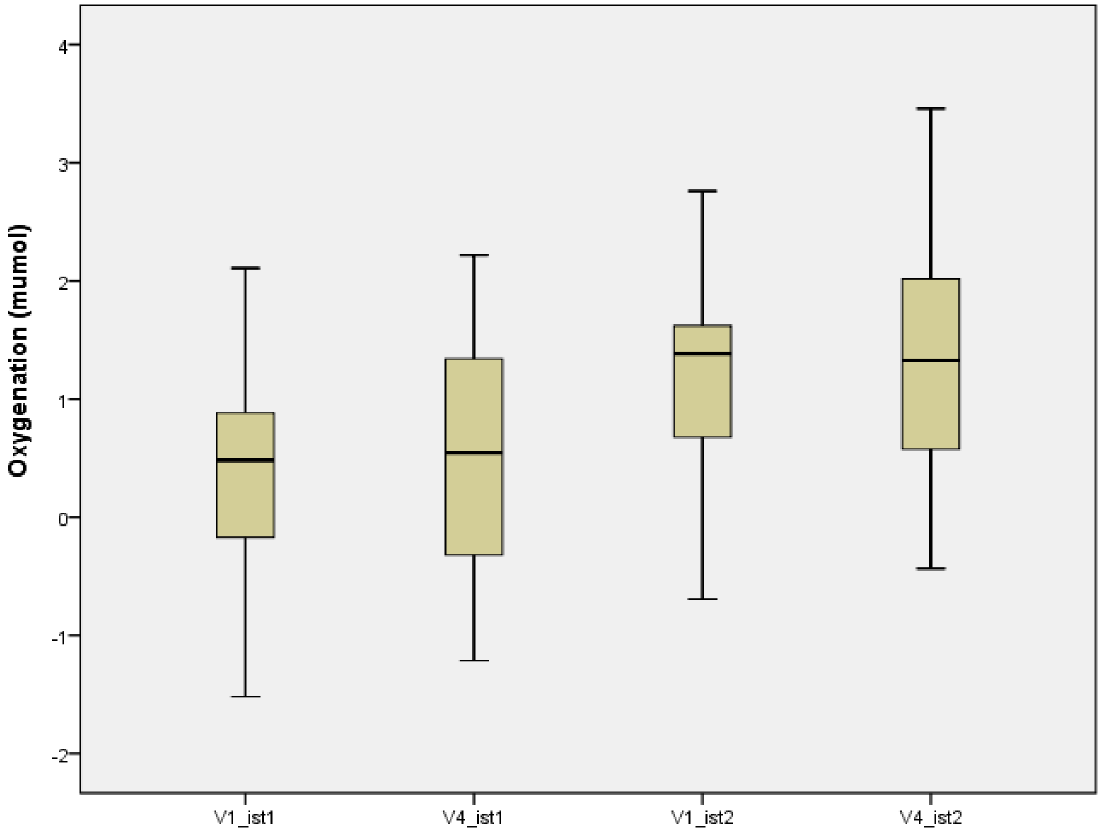

Figure 30: Boxplot of IST fNIR Measures

APPENDIX O

FNIR MEASURES DURING N-BACK TASK

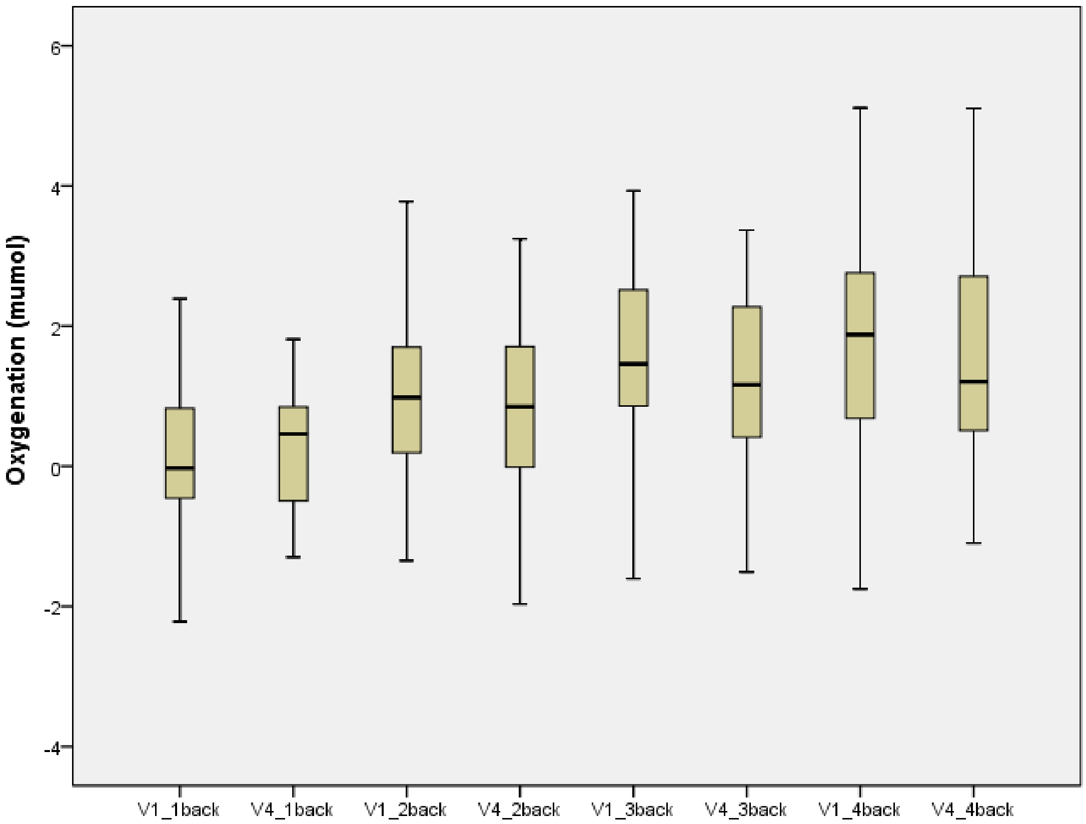

Figure 31: Bocplot of N-back fNIR Measures

APPENDIX P

THE PERSONALITY TRAIT STATISTICS

Table 60: Descriptive Statistics of Personality Traits

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Extraversion	37	1.50	4.88	3.52	0.73
Conscientiousness	37	1.50	4.25	3.18	0.67
Agreeableness	37	2.63	5.00	4.14	0.53
Neuroticism	37	1.56	3.89	2.64	0.64
Openness	37	2.17	5.00	3.60	0.64
Negative valence	37	1.00	2.83	1.59	0.41
Valid N (Listwise)	37				

APPENDIX R

THE PERSONALITY TRAIT RESULTS

Table 61: Detailed Results of the Personality Trait

Pr. No.	PF1	PF2	PF3	PF4	PF5	PF6
1	4.5	1.75	4.88	3.11	4.5	1.33
2	2	2.13	3.5	2.78	3	1.83
3	3	3.13	4.13	2.78	3.33	1.5
4	2.75	3.38	3.88	2.11	3	1.17
5	4.13	1.5	3.63	2.22	4	2.83
6	3.75	3.25	4	2.78	3.17	2.17
7	3	2.25	4.25	1.89	4	1.5
8	1.5	2.38	4.75	3.11	2.67	1.5
9	3.13	4	3.88	1.56	4	2
10	3.25	3.75	4	2	3.17	1.5
11	4.13	3.25	2.63	3.89	3.17	1.5
12	3.63	4.13	2.75	3.33	2.67	1.33
13	3	3.13	4.5	2.44	4.33	1.33
14	3.13	3.38	4.13	2.56	3.5	1.67
15	4.25	3.5	5	1.89	4.17	1
16	4.63	3.88	4.63	2.56	4	1.17
17	4.63	3.75	4.88	3.22	5	1.67
18	3.63	3.13	4.38	2.11	4	1.17
19	3.38	2.38	4	3.89	3.17	2.5
20	3.88	4.25	4	2.89	3.67	1.5
21	3.25	3.75	4.13	2.56	2.17	2
22	3.13	3.88	4.13	3.22	3.17	1.83
23	3	2.25	4.75	1.78	3	1.17
24	3.75	2.88	3.63	2.67	3.67	1.67
25	4	3.38	4.25	1.67	4	1
26	4.25	3.75	4.25	2.33	4.5	1.83
27	3.75	3.25	3.5	3.56	4.67	2.33
28	3.5	3.63	4	3	3.5	1.33
29	3.5	3.63	4.25	2.22	4	1.33
30	3.25	3.38	4.25	3	3.17	1.33
31	3.75	3.75	4.13	3.11	4	2
32	4.88	3.25	5	1.78	3.67	1.33
33	2.75	3.38	4	2.33	3.33	1.17
34	3.5	2.63	3.88	1.67	3.33	1.5
35	2.38	3.38	4.63	3.78	2.5	1.5
36	3.75	2.25	4.38	3	4	1.5
37	4.5	3.13	4.25	3	4	1.67

APPENDIX S

THE PERSONALITY TRAIT HIGHER/LOWER RESULTS

Table 62: Higher/Lower Results of the Personality Trait

Pr. No	PF1	PF2	PF3	PF4	PF5	PF6
1	higher	lower	higher	higher	higher	lower
2	lower	lower	lower	higher	lower	higher
3	lower	lower	lower	higher	lower	lower
4	lower	lower	lower	lower	lower	lower
5	higher	lower	lower	lower	higher	higher
6	higher	lower	lower	higher	lower	higher
7	lower	lower	higher	lower	higher	lower
8	lower	lower	higher	higher	lower	lower
9	lower	higher	lower	lower	higher	higher
10	lower	higher	lower	lower	lower	lower
11	higher	lower	lower	higher	lower	lower
12	higher	higher	lower	higher	lower	lower
13	lower	lower	higher	lower	higher	lower
14	lower	lower	lower	lower	lower	lower
15	higher	higher	higher	lower	higher	lower
16	higher	higher	higher	lower	higher	lower
17	higher	higher	higher	higher	higher	lower
18	higher	lower	higher	lower	higher	lower
19	lower	lower	lower	higher	lower	higher
20	higher	higher	lower	higher	higher	lower
21	lower	higher	lower	lower	lower	higher
22	lower	higher	lower	higher	lower	higher
23	lower	lower	higher	lower	lower	lower
24	higher	lower	lower	lower	higher	lower
25	higher	lower	higher	lower	higher	lower
26	higher	higher	higher	lower	higher	higher
27	higher	lower	lower	higher	higher	higher
28	higher	higher	lower	higher	lower	lower
29	higher	higher	higher	lower	higher	lower
30	lower	lower	higher	higher	lower	lower
31	higher	higher	lower	higher	higher	higher
32	higher	lower	higher	lower	higher	lower
33	lower	lower	lower	lower	lower	lower
34	higher	lower	lower	lower	lower	lower
35	lower	lower	higher	higher	lower	lower
36	higher	lower	higher	higher	higher	lower
37	higher	lower	higher	higher	higher	lower